

nature

THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE

*Rethinking the
electron microscope
as a tool for making
new materials*

PAGE 485



ATOM-UP CONSTRUCTION

GENETICS

A BIRD WITH FOUR SEXES

*The peculiar chromosomes of
the white-throated sparrow*

PAGE 482

RENEWABLE ENERGY

LET'S PRINT SOLAR CELLS

*Thin-film photovoltaics
should inseat silicon*

PAGE 488

PLANT DISEASES

TURNING OFF THE WATER

*Host and bacterial pathogen
fight for survival*

PAGES 506 & 524

NATURE.COM/NATURE

24 November 2016 \$10

Vol. 539, No. 7630



THIS WEEK

EDITORIALS

PHYSICS New light source can plant seeds of hope **p.468**

WORLD VIEW Europe set to act on pesticide chemicals **p.469**



IN ORBIT New GPS satellites to boost abilities on Earth **p.473**

Prevention is as good as a cure

Priorities for the US Cancer Moonshot Initiative face an uncertain funding future — but it must not ignore proven prevention programmes in favour of glitzy research.

As the United States and the wider world tries to work out what President Donald Trump will and won't do, one thing is clear: some Republican lawmakers feel no compulsion to govern until he takes office in January. This could leave funding for some of President Barack Obama's scientific mega-initiatives — including the Precision Medicine Initiative and the Cancer Moonshot Initiative — in limbo until next year. At present, there is a push in Congress to delay any decisions on the budget and pass a continuing resolution that will hold funding levels steady until March.

The Cancer Moonshot Initiative, one of the newest schemes on the block, had been scrambling to get its goals in order ahead of the budget decision in time to impress lawmakers with its worthiness. But it could face more months of uncertainty — a harsh outcome for an initiative that had already committed to an aggressive timeline of doubling the pace of cancer research in five years.

In recent months, the moonshot's architects have laid out an ambitious menu of projects, and prioritizing those was always going to be a challenge. Given the absence of a clear fiscal future, deciding which projects take priority has become even more crucial.

How to choose? It may be tempting to select and focus on the most superficially impressive and eye-catching science. But one of the moonshot's greatest strengths has been its willingness to take on not only cutting-edge research, but also the more mundane hurdles that can block the translation of findings to the clinic. Those low-tech projects should continue to receive emphasis, even as project leaders grapple with financial reality.

When the moonshot's advisory panel of cancer experts released its ten scientific recommendations in September, for example, it highlighted the expanded use of proven cancer-prevention techniques. In doing so, it drew welcome attention to these techniques and how improving their deployment could reduce deaths — from cervical cancer by 90%, from colorectal cancer by as much as 70% and from lung cancer by up to 95%.

The tools are at our fingertips, if only society could muster the means to use them. Take smoking — it is still on the rise worldwide, despite its clear link to cancer. In 2015, about 30% of US cancer deaths were from lung cancer. By 2030, smoking is projected to kill 8 million people globally each year. Researchers need to learn more about the barriers that have held back efforts to reduce these numbers, and how best to implement strategies for stopping smoking.

Similarly, colorectal cancer could be reduced through wider use of recent advances in non-invasive screening methods. But not many physicians are aware of these techniques, and too few people who are at risk of the disease are targeted for early screening.

Many cancers caused by human papillomavirus (HPV) are now preventable with a vaccine given to adolescents. However, uptake of the vaccine still falls below targets, in part because of parental concerns and misconceptions about the risks of vaccination.

Following the September report, the Moonshot Task Force — a separate group looking at how government agencies can work together to accelerate cancer research — released its own recommendations. These, too, included unglamorous but vital initiatives, such as a push for stop-smoking strategies in people who receive Medicaid, a US health-care plan for those with limited resources. And it announced

“When budgets are constrained, it is important not to let the glitzy outshine the familiar.”

an offer from some car-sharing companies to give discounts to cancer patients travelling to and from doctor's appointments. This helps to tackle one of many logistical hurdles that contribute to low public participation in clinical trials.

The more high-tech recommendations from the Moonshot Initiative are also worthy projects. They include large-scale genetic analysis of tumours to push forward the fledgling field of personalized medicine, and clinical-trial networks to better harness the information gleaned from trials of immunotherapies and treatments for paediatric cancers.

When budgets are constrained, it is important not to let the glitzy outshine the familiar. Too often in health research, straightforward prevention programmes, and the social science needed to implement them correctly, are pushed aside in favour of the basic research that fills this journal. A balance must be struck: the Cancer Moonshot and other initiatives should continue to emphasize the uptake of proven prevention strategies, even if the next administration tightens funds further than expected. ■

Daunting data

The power of big data must be harnessed for medical progress. But how?

There is art in 'big data' — in the poetic claims that it competes in volume with all the stars in the firmament. And in the seductive potential of its exponential, uncontrolled, ungraspable growth to improve our lives: by allowing medical treatments to be developed and approved more quickly — and, ultimately, truly personal medicine.

But at a workshop held in London by the European Medicines Agency earlier this month, just how much science has to happen to make this beautiful future a reality was apparent to all. Patient groups and research scientists attended, alongside computational heavyweights from IBM Watson Health and Google Cloud Platform. Together, they tackled chewy questions to which there are few answers.

How many data are 'enough' to reliably predict clinical effect? Which

data sets can be useful? How can they be managed? What's the best way to win the confidence of public and regulators? And, crucially, is academia training enough mathematicians and medical-data scientists, who will have to develop and harness all this new potential? The last of these questions at least has a clear answer: no.

Big data sets in medicine include genomics, transcriptomics and proteomics (which, respectively, describe our genomes, identify which of our genes are being expressed and catalogue all of the proteins in a tissue sample). Genomic data sets alone have already shown their value. The presence or absence of a particular gene variant can put people in high- or low-risk groups for various diseases and identify in some cases which people with cancer are likely to respond to certain drugs.

But a single molecular data set will not contain sufficient information to tell the whole story of an individual's medical fate. Integration of different types of molecular data might tell more, but that remains a computational challenge. Even more would emerge if an individual's molecular data were placed in the context of their physiology, behaviour and health. Electronic health records, the numbers of which are skyrocketing, could be useful here. So could disease registries, hospital and health-insurance records, as well as research publications and clinical-trial data. New types of data set come from wearables and apps, which collect health data directly from individuals, and from genome sequences of volunteers. Some researchers are even trying to extract medically relevant data from social-media platforms such as Twitter.

This adds up to a mind-boggling volume of data. According to one estimate at the workshop, clinical data from a single individual will generate 0.4 terabytes of information per lifetime, genomics data around 6 terabytes and additional data, 1,100 terabytes. By 2020, the amount of health-related data gathered in total will double every 73 days. Health professionals will confront more data than do those in finance.

To collect and hold all of this information within strict privacy regulations — non-negotiable for medical data — is another challenge.

Some tech firms, such as IBM Watson Health and Hewlett-Packard, are building systems that keep data local — algorithms can dip into them, but the data are not transmitted anywhere else. Google, unsurprisingly, thinks that all these data are safer on the Cloud.

The big question for scientists is how to take the next step to convert these artistic sketches of potential into scientific knowledge. Data sets vary in their reliability. Those derived from skimming social media will be very messy and will have to prove their usefulness. And large data sets, reliable or not, inevitably throw up spurious correlations,

“Health professionals will confront more data than do those in finance.”

and so recognizing meaningful patterns requires deep understanding of biology, something that software developers don't generally have. Future biologists — funders and universities should note — will need much more training in mathematics and data science.

Big data has the exciting potential to allow clinical trials to be conducted partly *in silico* — which would mean using fewer animals in drug testing as well as recruiting fewer patients to the actual trials. Yet the field is developing at a time when public trust in experts is at an all-time low.

Regulators are ready to be persuaded to accept computational information in clinical trials. On occasion, the European Medicines Agency and the US Food and Drug Administration already accept it in pharmacokinetics, one of the simplest data sets that drug developers must supply agencies with. But the trust of doctors, patients and regulators in the abstract informatics and mathematical know-how that go into developing scientific and clinical predictions cannot be taken for granted. As medics and researchers nudge the field gradually from poetry and seduction to delivery, they must engage the next generation of scientists — and the public — and inform and educate them in the art and science of what is possible. ■

Source of hope

More funding for the Middle East X-ray factory could give it a cafe — and sow the seeds of peace.

Even as the United States bombed Vietnam and the Red Army quashed an uprising in Czechoslovakia, there was a place where citizens from the two cold-war superpowers could meet: the CERN cafeteria. To this day, the same spot at Europe's particle-physics laboratory near Geneva, Switzerland, is a rare venue for Pakistani and Indian physicists to have a coffee, or for engineers from Israel and Iran to share ideas — no doubt fuelled by the huge lunches and inspiring views.

There is no better cure for cross-border acrimony than meeting someone from the other country and realizing that they are not so different from yourself. And major science facilities, whose inhabitants already share many beliefs, such as in openness and the scientific method, are natural places to kindle such mutual understanding.

What a chance, then, for SESAME. Synchrotron-light for Experimental Science and Applications in the Middle East is a venture that brings together Bahrain, Cyprus, Egypt, Iran, Israel, Jordan, Pakistan, the Palestinian Authority and Turkey, to build the region's first synchrotron-light facility near Amman in Jordan. Proposed in 1997, modelled on CERN and created under the auspices of the United Nations Educational, Scientific and Cultural Organization (UNESCO) in 2002, the remarkable collaboration is now set to open (see page 475).

SESAME has struggled at every turn. Several member states have made only partial payments, and its opening has been delayed by more than five years owing to financial difficulties. Other hurdles were less predictable: 2009 and 2010 saw Iranian council members assassinated,

and in 2013, the facility's roof fell in after unprecedented snowfall.

But SESAME has also ridden on a wave of generosity and perseverance. Its director and various advisers contribute their time for free. Facilities in Germany, France, the United Kingdom, Sweden, Spain and Switzerland have donated crucial parts of the machine, and Jordan, which provided land for the site at no cost, has continued to cough up funding, even as the neighbouring Syrian war has brought huge numbers of refugees across its borders.

To get this far is a terrific success. But when it opens next year, the facility will have just two beamlines, rather than a planned four, and will be without an administration building, dormitory, major library, auditorium or visitor centre. Nor, sadly, will SESAME have its hub of scientific life, a cafeteria. Experimental equipment takes priority, but in a facility where time is so precious that researchers allotted 24 hours of beam time barely sleep, some worry that, with nowhere to grab a coffee and a seat, scientists from conflicting countries may pass like ships in the night.

A group from the Sharing Knowledge Foundation in Geneva is trying to raise €30,000 (US\$32,000) through public donations by the end of this year towards creating a SESAME cafeteria. But funds for both the cafe and the science could also come from another source, which has been conspicuously lacking over the years: the United States.

A Congress ruling in 2011 cut funding to UNESCO, after the organization formally recognized Palestine as a member state. This is a public-relations difficulty for SESAME, even though there should be no official block because it is independent from UNESCO. A bigger question could be whether the incoming US president meddles with the Iran nuclear deal, which could restore sanctions.

Still, brokering peace in the Middle East is something that Donald Trump said on the campaign trail he'd “really like to do”. For less than the cost of a return flight to Tel Aviv on Air Force One, his administration could help SESAME to develop from a bare-bones scientific laboratory into a true beacon of peace and cooperation. ■



Stand firm on hormone disruptors

Ahead of a key meeting on endocrine-disrupting chemicals, Leonardo Trasande argues that policy must follow the science.

As the election of Donald Trump promises a bonfire of environmental regulation in the United States, Europe is poised to take a significant and possibly decisive step on how to regulate endocrine-disrupting chemicals (EDCs). These chemicals are everywhere — in food, personal-care products, electronics and furniture — and are widely detected in human blood and urine at levels known to affect health. Yet action on them lags behind controls on hazards such as carcinogens. Early next month, European Union member states will take an important step when they review criteria proposed by the European Commission for identifying and regulating EDCs.

Many pesticides contaminate foods and disrupt hormonal functions that are critical for brain development. Once EDC criteria are formalized, removal of these pesticides could help to prevent autism and loss of cognition, which have been linked to exposures *in utero*. The EDC criteria will also set scientific precedents for other national and global chemical policies.

The state of the science and policy discussions on EDCs are similar to those around climate change a decade ago. Research has suggested a more than 99% probability that these chemicals contribute to disease and disability. International bodies, including the World Health Organization (WHO), the United Nations Environment Programme and the International Council on Chemical Management, list them as an emerging public-health concern. The effects of EDCs cross the entire lifespan, with disease burden and costs of US\$217 billion annually in Europe and \$340 billion in the United States (T. M. Attina *et al. Lancet Diabetes Endocrinol.* <http://doi.org/bs55>; 2016). Even the reinsurance industry has advised its clients to reduce financial exposure related to the manufacture and use of EDCs.

In response to this evidence, lobbyists and environmentalists have traded rhetoric and warnings. And as with the debate on climate change, a small group of scientists — many with well-documented links to industry — have endeavoured to manufacture a level of doubt that is out of proportion to the level of scientific disagreement.

The scientists who deny endocrine disruption and dismiss the expert reviews on EDCs make many scientific inaccuracies and misrepresentations. Critics dismiss low-dose, nonlinear and non-monotonic exposure-response relationships for EDCs, even though they are well documented. They select studies with contaminated controls and other methodological problems to claim limited effects. They have argued that many studies of EDCs are based on correlation, not causation.

Endocrine-mediated adverse outcomes are complex. Too often lost in the debate, however, is that findings about the impacts of EDCs carefully control for confounding factors. Results in humans are consistent with those from the laboratory, strengthening the evidence for causation.

What should be discussed when the criteria are reviewed next month?

The WHO defines EDCs as “exogenous compounds or mixtures that alter function(s) of the endocrine system and consequently cause adverse effects in an intact organism, or its progeny, or (sub)populations”. With a seemingly innocuous edit, the European Commission’s draft criteria change the word “consequently” to “are known to” — placing too heavy a burden of proof for a chemical to be classified as an EDC.

Unlike carcinogens, mutagens and reproductive toxicants, which can be identified from animal studies under EU law, the draft EDC criteria require human data on health effects. Given that these can arise years (if not decades) after exposures in early life, an entire generation could suffer the health consequences of a regulatory delay. Animal and laboratory studies should be admissible. Arguably, this is the most important change needed.

Another major error by the commission is to misrepresent the crucial distinction between hazard and risk under EU law. The distinction means that considerations of potency — the traditional exposure-response effect — should not enter into the decision on whether, say, a pesticide exposure presents a hazard. This important principle is crucial to underwriting the regulation of endocrine disruptors, and was sealed as scientific consensus at an April meeting in Berlin. This meeting united the vocal minority of scientists with the leading authors of a scientific statement by the Endocrine Society (A. C. Gore *et al. Endocr. Rev.* <http://doi.org/bs69>; 2015). Participants agreed that the WHO definition was adequate, and that potency is not relevant to the identification of hazards such as EDCs.

The European Parliament should use science-based criteria to protect human health. EDC criteria should acknowledge the evolving weight of evidence of a chemical’s disruption of hormones and its contribution to adverse outcomes. This approach would allow chemicals to be designated as EDCs; suspected EDCs; endocrine-active substances; and endocrine-inactive substances. Like the approach used to re-evaluate potential carcinogens and reproductive toxicants, the designation can be reconsidered as new evidence emerges.

Some authors of the Berlin consensus statement continue to argue against the need for public-health protections, and have resorted to personal attacks, labelling peer-reviewed academic research as pseudoscience. In contrast to the US media, which has been criticized for its difficulty in discerning fact from fiction in the election campaign, we must stand firmly to defend scientific norms. The alternative is that public mistrust makes anti-scientific alternatives acceptable. ■

Leonardo Trasande is associate professor of paediatrics at New York University School of Medicine, USA.
e-mail: leonardo.trasande@nyumc.org

THE
EUROPEAN
PARLIAMENT
SHOULD PROTECT
HUMAN
HEALTH.

RESEARCH HIGHLIGHTS

Selections from the
scientific literature

ASTROPHYSICS

Homing in on a fast radio burst

The origins of powerful, millisecond-long radio pulses from space called fast radio bursts (FRBs) remain a mystery. But researchers studying the brightest FRB seen so far have zeroed in on its location more accurately than ever before.

Vikram Ravi at the California Institute of Technology in Pasadena, Ryan Shannon at CSIRO Astronomy and Space Science in Marsfield, Australia, and their team observed an FRB in real time using the Parkes Observatory radio telescope in Australia. By modelling the FRB's appearance in two of the telescope's 13 overlapping beams, the team constrained the origin of the FRB to just a handful of possible locations more than 1.5 billion light years (0.46 billion parsecs) away.

Pinpointing a burst's location should reveal its distance and allow astronomers to learn about the matter between galaxies, say the authors.

Science <http://doi.org/bs66> (2016)

PLANT BIOLOGY

Rapid recovery boosts plant yields

In strong sunlight, plants give off excess energy as heat to prevent damage, but when the light dims they are slow to switch off this process, reducing photosynthesis. Now researchers have shown that increasing the pace at which plants resume full photosynthesis can boost their yields by about 15%.

Stephen Long of the University of Illinois in Urbana, Krishna Niyogi of the University of California,

Berkeley, and their colleagues engineered tobacco plants (*Nicotiana tabacum*) by increasing the expression of three genes involved in protecting plants from bright light — *PsbS*, *VDE* and *ZEP*. They found that photosynthesis recovered more quickly under fluctuating light conditions than in unengineered plants. The dry weight of engineered plants grown in the field was 14–20% higher than that of normal plants.

This could be a way to increase the productivity of food crops, the authors say. **Science** 354, 857–861 (2016)

BIOCHEMISTRY

Catalyst clicks drugs together

A new type of catalyst assembles cancer drugs from benign components, and can work inside laboratory animals.

Some drug molecules can be split into smaller parts that are then equipped with specific chemical groups. These groups allow the components to be put back together in a 'click' reaction in the presence of a catalyst. This could reduce the side effects of cancer drugs by

genetic patterns, but the eDNA included extra sequences — suggesting that eDNA may reveal greater genetic diversity. Comparing the eDNA data with those from whale-shark groups worldwide showed that the Qatari animals are similar to other Indo-Pacific groups, which are genetically distinct from Atlantic populations.

Using eDNA is a powerful way to study population dynamics in marine animals, the authors say.

Nature Ecol. Evol. 1, 0004 (2016)



REINHARD DIRSCHERL/GETTY

BIOCHEMISTRY

A fast way to convert carbon

Researchers have come up with a biochemical pathway that converts carbon dioxide into useful organic compounds more efficiently than plants can.

A team led by Tobias Erb at the Max Planck Institute for Terrestrial Microbiology in Marburg, Germany, assembled a series of chemical reactions using 17 redesigned enzymes from 9 different organisms. The main enzyme used by plants in photosynthesis makes the process of CO₂ fixation relatively slow, but using other enzymes from bacteria, plants and humans, the researchers developed a quicker and more energy-efficient pathway.

This process could have many applications, such as boosting CO₂ fixation in photosynthetic organisms, the authors say.

Science 354, 900–904 (2016)

MICROBIOME

Low-fibre diet puts gut at risk

Mice eating a low-fibre diet have a higher risk of bowel infection, thanks to bacteria that normally live in the gut.

Eric Martens at the University of Michigan Medical School in Ann Arbor and Mahesh Desai, now at the Luxembourg Institute of Health, together with their colleagues, compared the effects of fibre-poor and fibre-rich diets in mice that lacked their own bacteria and were given a mix of 14 species of human gut bacteria. These microbes normally consume carbohydrates from dietary fibre, but without these nutrients, the bacteria instead degraded the mucus barrier that lines the intestinal wall.

Thinning of this protective layer (pictured as indicated by white arrows in the right panel; left panel shows normal layer) exposed the

intestinal surface to attack by disease-causing bacteria. More than half of the mice eating a low-fibre diet lost at least 20% of their body weight after infection with the pathogen *Citrobacter rodentium*.

Cell 167, 1339–1353 (2016)

GENOMICS

DNA reveals history of maize

Maize (corn) that grew more than 5,000 years ago was genetically more similar to today's maize than to its wild counterpart, suggesting that the plant was already being domesticated.

Nathan Wales and Jazmín Ramos-Madrigal of the University of Copenhagen and their team sequenced the genome of a 5,310-year-old maize cob (*Zea mays*) excavated in the 1960s in Mexico. They found a gene variant responsible for producing kernels that lack a hard seed coat and are easy to eat — probably the first step in domestication. However, other genes typically found in today's maize were absent, including one that prevents maize ears from shattering when ripe.

Curr. Biol. <http://doi.org/btcs> (2016)

DEVELOPMENTAL BIOLOGY

'Mini-guts' made with nerves

Miniature tissue structures resembling a human intestine with functional nerves can mimic the involuntary contractions that push food through the gut.

James Wells and Michael Helmrath of the Cincinnati Children's Hospital Medical

Center in Ohio and their team generated nerve and intestinal cells from human stem cells. They grew the mix for 28 days to make 3D structures, or organoids, that had intrinsic nerve cells. They then transplanted the organoids into mice to form mature nerves. A single electric impulse triggered a prolonged wave of contractions in the organoids that had a nervous system, but did not do so in structures lacking nerves. Using the model, the group identified molecular pathways involved in Hirschsprung's disease, a rare disorder of the large intestine.

These organoids could be used to study other intestinal diseases, the authors say.

Nature Med. <http://dx.doi.org/10.1038/nm.4233> (2016)

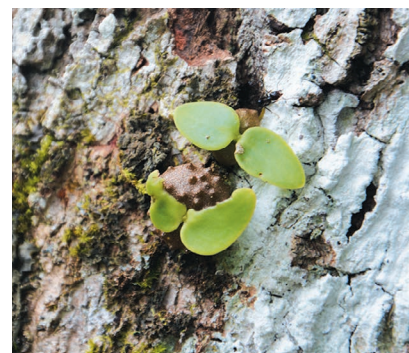
ECOLOGY

Ants farm plants in Fiji

A species of ant in Fiji survives by farming six plant species. The ants nurture seedlings and then live in the cavities of the plants.

Guillaume Chomicki and Susanne Renner of the University of Munich in Germany found that *Pholidris nagasau* ants carry seeds of six species of *Squamellaria* plant and insert them into cracks in trees, where they germinate (pictured). The ants also fertilize the seeds with their waste, and are only found living near these plants.

Although ants are known to farm fungi, this is the first time they have been found to plant seeds and actively cultivate them. Examining the family trees of relatives of these ants and plants suggests that this mutually dependent



relationship evolved around 3 million years ago.

Nature Plants <http://dx.doi.org/10.1038/nplants.2016.181> (2016)

GENOMICS

Epidemics marked indigenous DNA

Epidemics from Europe that killed thousands of indigenous Canadians in the nineteenth century have left their signatures in the genomes of people in Canada today.

Ripan Malhi at the University of Illinois in Urbana–Champaign, Michael DeGiorgio at Pennsylvania State University in University Park and their colleagues sequenced the genomes of 25 modern Tsimshian people, who live in coastal British Columbia and Alaska and are among Canada's First Nations. The authors compared the genomes with the DNA of 25 people who lived in the same region between 6,000 and 1,000 years ago. They found that most of the modern Tsimshian people had different versions of immune-system genes, including an HLA gene, from their ancestors. Moreover, the team found an overall decrease in genetic diversity, which probably signifies a sudden population decline of about 57% around 175 years ago.

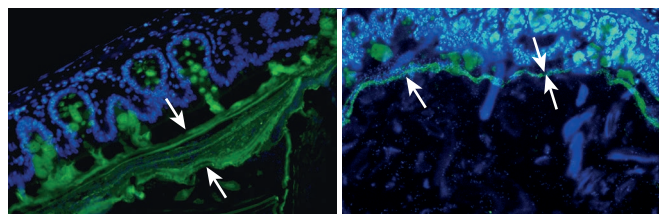
This is close to historical accounts that suggested a 70% decrease due to a smallpox epidemic.

Nature Commun. 7, 13175 (2016)

► **NATURE.COM**

For the latest research published by Nature visit:

www.nature.com/latestresearch



SEVEN DAYS

The news in brief

HEALTH

Zika virus trial

A trial that would release genetically modified mosquitoes in Florida to fight diseases such as the Zika virus can go ahead, the Florida Keys Mosquito Control District decided on 19 November. The initiative had been among those voted on by the state in non-binding referenda in the 8 November US election. The town of Key Haven, the proposed location of the trial, rejected the measure, but voters elsewhere in the broader area of Monroe County approved it. The mosquito board has the ultimate say. The experiment, which would be a US first, needs a new site and approval from the US Food and Drug Administration before it can begin. Separately, the World Health Organization announced on 18 November that the Zika virus no longer presents a health emergency of global concern.

Malaria vaccine

Vaccinations against malaria will begin in sub-Saharan Africa in 2018, the World Health Organization announced on 17 November. A handful of pilot projects will use the first ever malaria vaccine — called RTS,S — against the world's deadliest malaria parasite, *Plasmodium falciparum*. The projects will test the logistics of administering a vaccine that must be given to children 4 times over 18 months, and that yields only partial protection.

CRISPR in humans

A Chinese team of researchers has become the first to inject a person with cells modified using the CRISPR–Cas9 gene-editing technique, with the aim of treating aggressive lung

cancer. The participant is part of a clinical trial at the West China Hospital in Chengdu. See page 479 for more.

POLICY

Brexit and science

An influential group of politicians has urged the UK government to guarantee that all European Union researchers living in the country can stay after it leaves the union. In a report on the impacts of Brexit on science, released on 18 November, the House of Commons science and technology select committee advised the government to set out a broader vision for science — including committing to a hefty raise in the country's science spending. The committee also said that the

department in charge of Brexit should move swiftly to appoint a chief scientific adviser, a position currently unfilled and unadvertised. Separately, Prime Minister Theresa May announced on 21 November that her government would invest more in research and development — providing increases worth £2 billion (US\$2.5 billion) per year by 2020 — to keep Britain at the “cutting edge”.

Women only

In a bid to increase the proportion of women in high-ranking positions in academia, the Royal Netherlands Academy of Arts and Sciences will hold two special elections in 2017 to recruit 16 additional female scholars and scientists. Nominations for the first round

of women-only elections, the results of which are to be announced in May, are open until 8 December. Currently only 13% of the academy's 556 members are women. Normal annual elections for 16 new members, male or female, will continue in parallel with the two extra rounds, the academy says.

Climate concern

Australia, Canada and the United States all risk breaking the climate pledges that they made in Paris a year ago, according to an analysis of past and present action by the world's largest national economies. The study by the London School of Economics and Political Science, published on 17 November at the United Nations climate meeting in



UJET IFANSATI/GETTY

Peat initiative targets climate and health

Peatlands cover only 3% of the global land surface, but contain about twice as much carbon as the entire biomass of the world's forests. A global initiative launched on 17 November at the United Nations climate meeting in Marrakesh, Morocco, seeks support from governments and international organizations to scale up protection

of the organic carbon stock and restoration of drained wetlands. Drained and burned peatlands are responsible for about 5% of carbon emissions from human activity. Peat forest fires last year in Indonesia (pictured), one of the founding members of the initiative, have been linked to 100,000 premature deaths in the region.

STEPHANE CORVAUX/ESA/GETTY

Marrakesh, Morocco, finds that Argentina, Saudi Arabia and Turkey are likewise behind on meeting their climate goals. Only France, Germany and the United Kingdom are curbing their domestic greenhouse-gas emissions in line with their Paris targets; Brazil, China and Italy look set to almost keep their pledges, the authors say.

EVENTS

Super satellite

The Geostationary Operational Environmental Satellite-R Series (GOES-R), the United States' most scientifically capable weather satellite yet, successfully launched on 19 November. GOES-R can take pictures every 30 seconds, which is 5 times faster than current GOES satellites can, and in 4 times higher resolution. The rapid imaging will allow it to track severe storms, wildfire smoke and volcanic ash as they spread. GOES-R should be operational within a year.

Quadruple launch

A quartet of Galileo satellites successfully launched on an Ariane 5 rocket (pictured) from French Guiana on 17 November, bringing Europe's civil global positioning system close to completion. The 18 Galileo satellites now in orbit should allow the system to enter into preliminary



service by the end of the year. By 2020, a full constellation of 24 satellites along 3 orbital planes — plus 2 spares per plane — are to routinely transmit their position and the time, allowing users worldwide to locate objects on Earth with extreme precision.

Plant pathogen

The deadly *fastidiosa* subspecies of the plant pathogen *Xylella fastidiosa* has been identified on the island of Mallorca, Spain. It is the first appearance of this subspecies in outside environments in Europe; in April, it had been discovered in oleanders in a greenhouse in Germany. The subspecies is different from that responsible for the recent devastation of olive groves in Puglia, Italy — but the same as that which wreaked havoc last century in grapevines in California, where *X. fastidiosa*

is endemic. The bacterium was first identified in Europe in 2013.

FACILITIES

ITER creeps on

The council of the nuclear-fusion experiment ITER endorsed a radical change of plans on 17 November: the international project will now aim to store hydrogen plasma in its doughnut-shaped container by 2025, postponing full nuclear fusion until 2035. The plan was proposed in 2015 by director-general Bernard Bigot in an attempt to save the seriously late and over-budget project. In spring 2016, two external reviews expressed cautious optimism about Bigot's management changes and the pace of construction at the site in St-Paul-lez-Durance, France.

AWARDS

John Maddox prize

Human-memory researcher Elizabeth Loftus has been awarded the 2016 John Maddox Prize. Loftus, at the University of California, Irvine, has worked on the cognitive psychology of false memories and unreliable eye-witness testimony. She has influenced the US legal system, suffered severe harassment in reaction to her work and highlighted the vulnerability of academics

COMING UP

28 NOVEMBER–
2 DECEMBER

Dark matter and gravitational waves are top of the agenda at the 13th International Symposium on Cosmology and Particle Astrophysics in Sydney, Australia.

go.nature.com/2gc0luo

1–2 DECEMBER

Delegates from Canada and the 22 member states of the European Space Agency meet in Lucerne, Switzerland, to discuss the agency's focus for the coming years.

go.nature.com/2f7wmuk

in the face of legal threats. The John Maddox Prize is awarded annually for championing sound science and evidence in the face of hostility and repression. It is sponsored by *Nature* and the London-based charities the Kohn Foundation and Sense About Science. It is named after the late John Maddox, a former editor of *Nature*. See go.nature.com/2gdrxxo for more.

ENERGY

Texas tea

A massive deposit of untapped hydrocarbon deposits, thought to be the largest ever in the United States, has been discovered in western Texas. The Midland Basin of the Wolfcamp Shale area contains an estimated 20 billion barrels of oil and 1.6 billion barrels of natural gas, the US Geological Survey said on 15 November. The agency has not assessed whether it would be profitable to produce the 'unconventional' resources, which would have to be extracted using special methods.

➔ NATURE.COM

For daily news updates see:

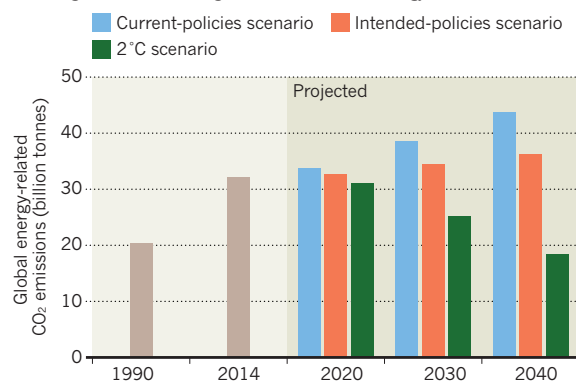
www.nature.com/news

TREND WATCH

The carbon budget to keep global warming below 2°C could be used up by 2040 under current climate policies, says the International Energy Agency in its latest *World Energy Outlook*. The report makes projections of global energy-related carbon dioxide emissions up to 2040 under three scenarios (see chart). Enforcing pledges in the 2015 Paris climate deal offers only a few years' respite. An early peak and subsequent decline in emissions in line with meeting the 2°C target would require extra efforts, the agency says.

CLIMATE PLEDGES FALL SHORT OF 2°C TARGET

Current and probable policies to curb carbon emissions will fail to limit warming to 2°C, according to the 2016 *World Energy Outlook*.



Current policies are those enacted as of mid-2016; intended scenario includes announced policies that will probably be enacted.

NEWS IN FOCUS

US ELECTION Minority scientists rethink careers in wake of Trump win **p.476**

BIOTECHNOLOGY CRISPR gene-editing moves into the clinic in China **p.479**

BRAZIL Researchers fight to avert catastrophic funding freeze **p.480**



BIOLOGY The bizarre chromosome evolution of a common bird **p.482**

ADAM HART-DAVIS/SPL



The SESAME accelerator will circulate electrons at high energies, creating beams of intense radiation that will be used to study cultural treasures.

PHYSICS

Middle Eastern X-ray factory readies for action

SESAME project set to revolutionize science in the region, but is strapped for cash.

BY ELIZABETH GIBNEY

The Middle East's first major international research centre has weathered political unrest, international sanctions and even the assassination of two delegates. Now, the Synchrotron-light for Experimental Science and Applications in the Middle East, or SESAME, is on the brink of circulating its first subatomic particles.

The machine, which lies outside Amman,

Jordan, will start accelerating electrons around its 133-metre ring in December and begin using the resulting beams of intense radiation to do science from May 2017.

Strapped for cash, SESAME will initially operate at half its planned capacity. But its location in the Middle East means that it will make a significant contribution to science. "It will set the stage for a culture of state-of-the-art technology and science in our region," says director Khaled Toukan.

SESAME is a collaboration between Bahrain, Cyprus, Egypt, Iran, Israel, Jordan, Pakistan, the Palestinian Authority and Turkey. Created under the auspices of the United Nations Educational, Scientific and Cultural Organization (UNESCO), it has also received funding from other countries and from the European Union.

The machine will generate light in a range of wavelengths and then channel it into intense 'beamlines' of specific frequencies (see 'Open SESAME'). These can be used to reveal the ►

OPEN SESAME

The SESAME synchrotron near Amman, Jordan, will start circulating beams in December. But it will initially use only a fraction of its potential capacity.

MICROTRON

Electrons accelerated to 22 megaelectronvolts.

BOOSTER RING

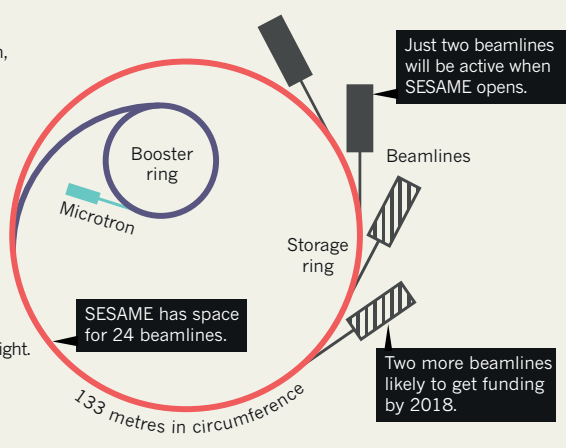
Electrons further accelerated to 800 megaelectronvolts.

STORAGE RING

Electrons circulate at 2.5 giga-electronvolts, and start to emit light.

BEAMLINES

Optical instruments tailor the light and use it to investigate samples.



► structure of materials, biological samples and artefacts down to the atomic scale. SESAME has space for 24 beamlines. The plan was initially to open with four of the slots filled, but a lack of funding means that it will open with just two — one infrared and one X-ray.

“SESAME is in a part of the world where you have very difficult times politically, and a lot of instability, and the money just didn’t come,” says Toukan. Iran couldn’t pay its contribution until January, owing to sanctions that prevented it from transferring money internationally. Cyprus has not paid its share of the running costs since it was hit by a financial crisis in 2011, says Toukan, and Pakistan has paid only half of its dues this year. Jordan and Turkey are the only countries to have paid their annual and capital contributions in full, he says. Politics has touched SESAME, too: two Iranian physicists killed in car-bomb attacks in 2009 and 2010 were members of its council.

Still, engineers are expected to have finished testing the synchrotron by May, and a group of 260 researchers, largely from universities in the Middle East and in fields from pharmacology

to physics, will then be able to apply for time on the two beamlines.

Toukan hopes that the facility will ease brain drain from the region, boost collaboration between the scientific community and industry and create opportunities for researchers who lack the funds to leave the Middle East.

Because of its proximity to many of the world’s archaeological treasures, the facility presents a fresh opportunity for studying the region’s cultural heritage. Jan Gunneweg, an archaeometrist at the Hebrew University of Jerusalem, hopes that scientists will use SESAME to collaborate on understanding their shared history. Many artefacts — such as Egyptian mummies and delicate papyrus — are fragile and must be insured at high cost if they are to travel long distances. “If that material has to go in the air, you destroy it,” says Gunneweg.

He wants to use the synchrotron to further his studies of the composition, and therefore origins, of parchments including the Dead Sea Scrolls, the oldest known biblical texts.

SESAME cost around US\$110 million to build, just one-sixth of the price of the European

Synchrotron Radiation Facility in Grenoble, France, one of the world’s most advanced synchrotrons. Pared-back ambition helped, but SESAME also relied on parts donated from dismantled European facilities, and was built on land that was given for free.

Two more beamlines are due to be installed by the end of 2018. SESAME has secured \$2 million from the Jordanian Scientific Research Fund for its third beam, which will perform protein crystallography. And Toukan is confident that the collaboration will find funds for a fourth beam dedicated to materials science.

Beyond these four beams, there is no set schedule for filling the remaining 20 slots. “A few tens of million of dollars could make this dream come true,” says Roy Beck, a biophysicist at Tel Aviv University in Israel and a committee member of the SESAME users’ group. He laments that more nations have not been willing to swallow national differences for the sake of science. Some Gulf countries will not take part because of Israel’s participation, he says, and the United States has made only a small contribution, which both Beck and Toukan attribute to political considerations (see page 468).

Advocates hope that SESAME will foster peace in the same way that CERN, Europe’s particle-physics laboratory, near Geneva, Switzerland, helped to heal the wounds of the Second World War and brought Soviet and Western scientists together at the height of the cold war.

But crucial factors in that success were CERN’s communal spaces, where scientists could share a coffee and get to know each other, says Beck. SESAME is scheduled to open without a cafeteria or dedicated accommodation, although a committee is trying to raise US\$32,000 in donations to create the former.

“I hope people from all round will understand that this is a true chance for people within the Middle East to join hands and talk about things that unite them,” says Beck. ■

SOCIETY

Immigrant and minority scientists shaken by Trump win

Worries include job prospects, discrimination — and safety.

BY HEIDI LEDFORD, SARA REARDON & RAMIN SKIBBA

As the US presidential election results rolled in, Naglaa Shoukry watched a door slam shut. An immunologist at the University of Montreal in Canada, she had been contemplating a move to the United

States in search of better research funding. But when Donald Trump clinched the presidency, she knew that she would probably not go.

Shoukry, a Muslim from Egypt, did post-doctoral research in Ohio and was there when the United States tightened security after the terrorist attacks of 11 September 2001. When Trump pledged to use “extreme

vetting” to determine which immigrants could enter the country, Shoukry recalled the humiliations her family experienced when travelling to see her in Ohio. “You have an interesting name,” a US border official once told her brother, Mohamed, before detaining him for extra security checks. Under Trump, Shoukry decided, it would surely be worse.

Shoukry is not the only scientist shaken by the result of the 8 November election, nor is she alone in reconsidering whether to work or study in the United States. Trump's campaign rhetoric was at times insulting to women, immigrants and under-represented minorities, and there are signs that his victory has further inflamed racial tensions. The Southern Poverty Law Center, a civil-rights advocacy group in Montgomery, Alabama, collected 437 reports of intimidation and harassment in the five days after the election — many of which explicitly referenced president-elect Trump. Fifteen per cent of the incidents reported to the centre took place at universities.

Such events have led several universities and scientific societies to reaffirm their commitment to diversity. "Take concrete steps to protect and advocate for colleagues and students who are particularly vulnerable right now," urges a letter signed by ten astronomers on the Astronomy in Color blog, which advocates for diversity in the field. "We must reiterate how absolutely essential it is to the core values of our community, and also to the well-being of our society and world, that all persons be treated with the dignity and respect they deserve," wrote Amy Gutmann, president of the University of Pennsylvania in Philadelphia, when black undergraduates there received racist e-mails after the election. "We all stand together in solidarity with our Black students."

Still, scientists around the country have reported harassment. Mónica Feliú-Mójer, a science communicator originally from Puerto Rico, is nervous about speaking Spanish in some public places after hearing a passer-by shout "Build that wall!" — referencing Trump's plan to build a wall along the US–Mexico border. The 12 November incident took place in San Diego, California, where Feliú-Mójer was attending the Society for Neuroscience annual meeting.

Patrick Freeman, an ecologist at the Carnegie Institution for Science in Stanford, California, was out driving on the day after the election when a person in a truck with a Trump bumper sticker started honking at him — perhaps because Freeman's car sports a sticker in support of Trump's Democratic opponent, Hillary Clinton, and of lesbian, gay, bisexual, transgender and queer people. The truck driver pulled up alongside Freeman and repeatedly pantomimed shooting him before speeding away.

For some scientists, the post-election atmosphere has been stifling. Nicole Cabrera Salazar, an astronomy graduate student at Georgia State University in Atlanta who was born in Chile, says she is mindful that her field is dominated by white males — and that 63% of white males who voted supported Trump. Cabrera Salazar was once outspoken in support of marginalized students, but she has begun to hold back for fear of a backlash. "It's a toxic environment for women of colour," she says.



Donald Trump has taken a hard-line stance on US immigration.

Some foreign researchers even question whether to travel to the United States for conferences. Shaaban Mousa, an anaesthesiologist at Charité University Hospital in Berlin, comes to the Society for Neuroscience meeting every year, but considered skipping it this year because of the election result. "I was not relaxed," he said last week at the meeting. "I came because I have my ticket." And Razi Nalim, an engineer at Indiana University–Purdue University in Indianapolis, worries that he'll have difficulty recruiting Muslim students and postdocs to work in his department.

"'Why are you still here?' a man shrieked. I think he meant: 'Why are you still here after Trump's election?'"

About 2% of scientists in the United States are Muslim immigrants, according to a survey of biologists and physicists led by Elaine Howard Ecklund, a sociologist at Rice University in Houston, Texas. And 64% of them say they have experienced religious discrimination — a higher proportion than for any other racial or religious group in Ecklund's study. The US Federal Bureau of Investigation's annual hate-crimes report suggests that number may grow: hate crimes against Muslims in the United States climbed by 67% last year.

But several Muslim scientists contacted by *Nature* still feel comfortable in the United States. Some, such as a husband and wife studying neuroscience at the Medical University of South Carolina in Charleston, were pleased

to see Trump win. "It's a good change for the US," says the husband, who asked not to be named over concerns with his career. He and his wife say that they have never experienced racism or Islamophobia during their time in South Carolina — a state where Trump won 55% of the vote — and they have been shocked by mass e-mails from their professors calling Trump supporters racist and hateful.

Shoukry says that some of her Muslim colleagues in the United States are troubled by the thought of raising children in an increasingly hostile environment. Islam Hussein, an Egyptian virologist who is a research scientist at the Massachusetts Institute of Technology in Cambridge, says his children were nervous about going to school the day after the election. But after a decade in the United States, Hussein wants to stay. His family has always felt welcome, he says, and his wife, who wears a hijab, has a flourishing career as a pre-school teacher. After the election, friends rushed to offer the family support. "This is the spirit of America that we should all keep embracing," Hussein told *Nature* on 14 November.

The next day, Hussein contacted *Nature* again: his wife had just been accosted at their local pharmacy. "Why are you still here?" a man shrieked at her, then hopped into his car and fled. "I think he meant: 'Why are you still here after Trump's election?'" says Hussein. "This is the first explicit act of discrimination we have encountered in ten years." ■

Additional reporting by Erika Check Hayden.

Asteroid crater probed

Drilling into the Mexican site of the dinosaur-killing impact reveals how enormous craters form and finds shattered rock where underground life could thrive.

BY ALEXANDRA WITZE

Drilling into ground zero of the asteroid impact that killed off the dinosaurs 66 million years ago has uncovered the origin of its mysterious ring of mountains.

The drill core penetrated a circle of mountains, known as a peak ring, in Mexico's buried Chicxulub crater. Only the largest impacts are powerful enough to create peak rings. Understanding how these mountains formed at the 200-kilometre-wide Chicxulub could help to reveal how cosmic collisions shaped other bodies, such as the Moon and Venus.

The work also shows how, despite killing the dinosaurs, the impact may have created an environment in which other life could thrive. The cosmic smash-up fractured rocks, opening up spaces and warm habitats for microbes to move in.

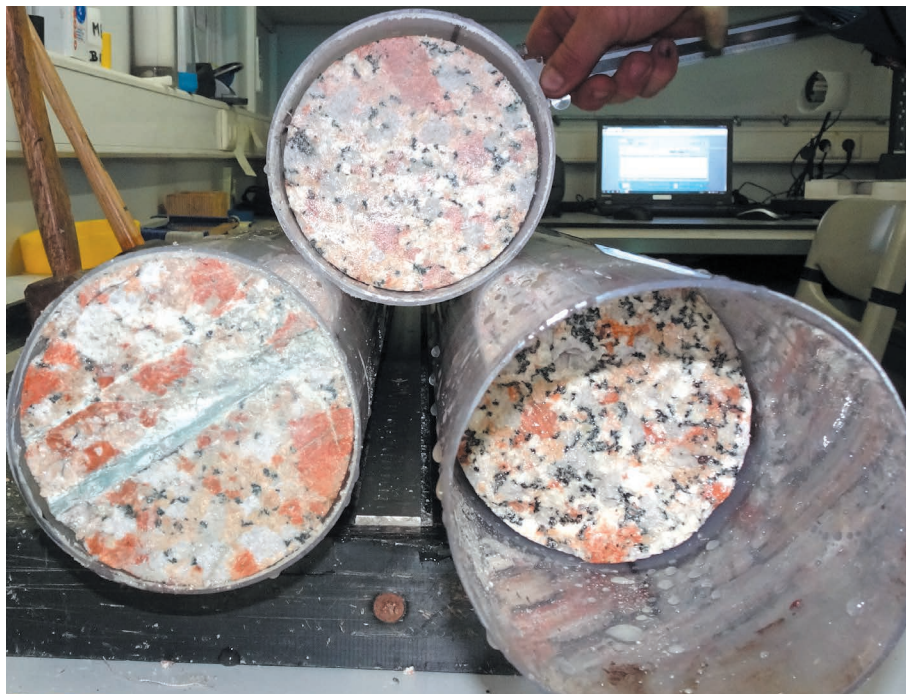
"We got a core better than we could reasonably have imagined," says David Kring, a geologist at the Lunar and Planetary Institute in Houston, Texas. "It is absolutely extraordinary." He and his colleagues described the findings on 18 November in *Science*¹.

After a decade of planning, the project penetrated 1,335 metres into the sea floor off the coast of Progreso, Mexico, in April and May. Drillers hit the first peak-ring rocks at a depth of 618 metres, and a pinkish granite at 748 metres.

Geologists know that the granite must have come from relatively deep in the crust — perhaps 8–10 kilometres down — because it contains big crystals. The size of these crystals suggests that they formed through the slow cooling of deep, molten rock; by contrast, rapid cooling at shallow depth tends to form small crystals. Finding the granite relatively high in the drill core means that something must have lifted it up and then thrown it down on top of other rocks.

That rules out one idea of how craters form, in which the pulverized rock stays mostly in place like hot soup in a bowl. Instead, the core confirms the 'dynamic collapse' model of cosmic impacts, in which the asteroid punches a deep hole in the crust, causing the rock to flow like a liquid and spurt skyward. That rock then falls back to Earth, splattering around in a peak ring.

"The ring of mountains we observe is made of deep material uplifted and flipped



A core from the Chicxulub crater contains pinkish granite that hints at how the basin formed.

on its head," says Sean Gulick, a geophysicist at the University of Texas at Austin. Gulick co-led the Chicxulub expedition with Joanna Morgan, a geophysicist at Imperial College London.

FREE SPACE

Knowing that Chicxulub formed through dynamic collapse opens new ways to explore other worlds, says Kring. Last month, he and his colleagues showed that, like Chicxulub, the peak ring of the Moon's 320-kilometre-wide Schrödinger crater is made of material from deep in the crust². This means that robots or astronauts could visit Schrödinger to pick up samples of the lunar interior.

Chicxulub is the only good example of a peak-ring crater on Earth. Studying peak rings in impact basins on other worlds could confirm whether the dynamic-collapse model holds true for different environments, says David Baker, a planetary scientist at NASA's Goddard Space Flight Center in Greenbelt, Maryland, who has published

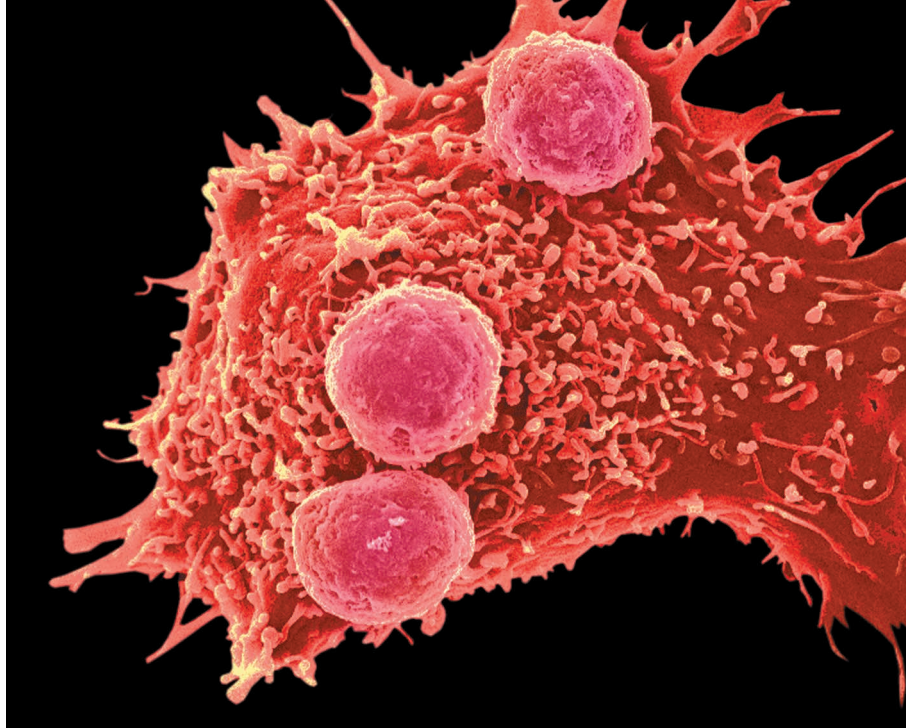
on peak-ring formation³.

The Chicxulub core also explains a long-standing geophysical mystery — why seismic waves move surprisingly slowly in the peak-ring rocks. The answer is that the impact fractured the granite so badly that it became much less dense than typical granite, slowing the waves' progress.

Lower density also means that rock is more porous, providing space for microbes to move around inside. After the sterilizing heat of the Chicxulub impact had dissipated, organisms may have taken advantage of the lingering warmth. "Somehow, life finds its way into these high-porosity targets and takes advantage of the habitat created," says Gulick. An underground biosphere could have evolved over millions of years, surviving off chemical energy and warmth without sunlight.

Project scientists have found cells and microbial DNA in the Chicxulub core, but are holding details for future publications. ■

1. Morgan, J. V. *et al. Science* **354**, 878–882 (2016).
2. Kring, D. A., Kramer, G. Y., Collins, G. S., Potter, R. W. K. & Chandnani, M. *Nature Commun.* **7**, 13161 (2016).
3. Baker, D. M. H., Head, J. W., Collins, G. S. & Potter, R. W. K. *Icarus* **273**, 146–163 (2016).



Gene editing could improve the ability of immune cells (spherical) to attack cancer.

BIOTECHNOLOGY

CRISPR gene editing tested in a person

Trial could spark biomedical duel between China and US.

BY DAVID CYRANOSKI

A Chinese group has become the first to inject a person with cells that contain genes edited using the revolutionary CRISPR–Cas9 technique.

On 28 October, a team led by oncologist Lu You at Sichuan University in Chengdu delivered the modified cells into a patient with aggressive lung cancer as part of a clinical trial at the West China Hospital, also in Chengdu.

Earlier clinical trials using cells edited with a different technique have shown promise at treating disease. The emergence of CRISPR, which is simpler and more efficient than other techniques, will probably accelerate the race to get gene-edited cells into the clinic, says Carl June, who specializes in immunotherapy at the University of Pennsylvania in

Philadelphia and led one of the earlier trials.

“I think this is going to trigger ‘Sputnik 2.0’, a biomedical duel on progress between China and the United States, which is important since competition usually improves the end product,” he says.

June is also the scientific adviser for a planned US trial that will use CRISPR to target three genes in cells extracted from participants, with the goal of treating various cancers. He expects the trial to start early next year. In March 2017, a group at Peking University in Beijing hopes to start three clinical trials using CRISPR against bladder, prostate and renal-cell cancers. Those trials do not yet have approval or funding.

Lu’s trial received ethical approval from a hospital review board in July. Injections into participants were supposed to begin in August

but the date was pushed back, Lu says, because culturing and amplifying the cells took longer than expected and then the team ran into China’s October holidays.

The researchers removed immune cells from the recipient’s blood and then disabled a gene in them using CRISPR–Cas9, which combines a DNA-cutting enzyme with a molecular guide that can be programmed to tell the enzyme precisely where to cut. The disabled gene codes for the protein PD-1, which normally puts the brakes on a cell’s immune response: cancers take advantage of that function to proliferate.

Lu’s team then cultured the edited cells, increasing their number, and injected them back into the patient, who has metastatic non-small-cell lung cancer. The hope is that, without PD-1, the edited cells will attack and defeat the cancer.

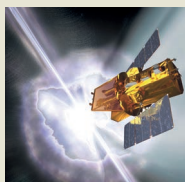
Lu says that the treatment went smoothly, and that the participant will get a second injection, but declined to give details because of patient confidentiality. The team plans to treat ten people in total; each will receive either two, three or four injections. It is primarily a safety trial, and participants will be monitored for six months to determine whether the injections are causing serious adverse effects. Lu’s team will also watch them beyond that time to see if they seem to be benefiting from the treatment.

Other oncologists are excited about CRISPR’s entry onto the cancer scene. “The technology to be able to do this is incredible,” says Naiyer Rizvi of Columbia University Medical Center in New York City. Antonio Russo of Palermo University in Italy notes that antibodies that neutralize PD-1 have successfully kept lung cancer in check, boding well for a CRISPR-enabled attack on the protein. “It’s an exciting strategy,” he says. “The rationale is strong.”

But Rizvi questions whether this particular trial will succeed. The process of extracting, genetically modifying and multiplying cells is “a huge undertaking and not very scalable”, he says. “Unless it shows a large gain in efficacy, it will be hard to justify moving forward.” He doubts it will be superior to the use of antibodies, which can be expanded to unlimited quantities in the clinic. Lu says that this question is being evaluated in the trial, but that it’s too early to say which approach is better. ■



TOP NEWS



Long-sought signal deepens mystery of fast radio bursts
go.nature.com/2fmo4df

MORE NEWS

- The maths of science’s broken reward system go.nature.com/2gxwq5w
- Laser used to control mouse’s brain — and speed up milkshake consumption go.nature.com/2gdcknz
- European diseases left mark on First Nations’ DNA go.nature.com/2fnrys2

NATURE PODCAST



Tracking whale shark DNA; human computers behind early astronomy; and building with a microscope nature.com/nature/podcast



Brazilian President Michel Temer wants to freeze real-terms federal spending for the next two decades.

POLITICS

Brazil's scientists fight funding freeze

Proposed law could restrict research spending for 20 years.

BY CLAUDIO ANGELO

Brazil's science budget has shrunk by more than 40% in the past 3 years. But researchers are now trying to escape an even worse crisis: federal science spending could be frozen at its current low level for the next two decades, under a constitutional amendment to cap public spending to inflation-level rises.

On 16 November, 19 institutions representing higher education and research sent a letter to Brazil's Senate asking that politicians exclude science and education from the amendment, which is President Michel Temer's top legislative priority as he aims to halt the country's spiralling fiscal crisis. The amendment passed the lower house of Brazil's National Congress last month; the Senate is due to vote on it by the end of this month.

If Temer's legislation does end up constricting federal science spending, "it will be the end of science in Brazil", warns physicist Luiz Davidovich, who is president of the Brazilian Academy of Sciences in Rio de Janeiro, which co-authored the letter.

Last month, scientists joined protests in Rio and other cities against the amendment, which they called the #PECdoFimDoMundo (amendment of the end of the world).

Brazil's public spending does need to be reined in, says Carlos Américo Pacheco, chief executive of the São Paulo Research

Foundation, a state science-funding agency. "It's a bitter medicine, but it's inevitable: you won't be able to fix public finance without a major effort to cut expenses," he says.

But scientists say that they've had enough pain. In 2013, the science-ministry budget stood at 7.9 billion reais (US\$2.3 billion). After

"It's a bitter medicine, but it's inevitable."

an economic slump, the budget has now fallen to 4.6 billion reais — the lowest level in 7 years, and the lowest in a decade when inflation is accounted for. And that includes the budget of the communications ministry, which Temer merged with the science ministry to create the Ministry of Science, Technology, Innovations and Communications (MCTIC) when he took office in May.

DRAMATIC MOVES

As a result of the cuts, some Brazilian scientists have left the country; others are making do on less funding. And officials at major planned facilities such as the 850-million-real Multipurpose Research Reactor outside São Paulo — which aims to generate medical radioisotopes — and the 510-million-real Sirius synchrotron, in Campinas, have complained that they haven't yet received their 2016 funds.

Freezing the science budget at these low levels could be disastrous, says Helena Nader,

president of the Brazilian Society for the Advancement of Science. Federal science spending is currently around 1.1% of Brazil's gross domestic product (GDP), but that figure will fall if spending is frozen and GDP grows over the coming years. Nader compares that with science spending in China, which is planned to grow from 2.1% of GDP now to 2.5% in 2020.

The MCTIC says that the amendment will not necessarily affect the science budget. Science might even get more money — provided that other areas of federal funding get less. But critics say that's unlikely in a country with other pressing needs, including chronically underfunded public education and health and a collapsing pensions system.

NEXT YEAR'S BUDGET

MCTIC minister Gilberto Kassab is trying to increase the science budget ahead of the amendment. His 2017 request, included in a government proposal that has not yet been voted on by Congress, would increase the budget by 28% to 5.9 billion reais.

"The MCTIC has managed to include in the annual budget a value that's enough to meet the needs of the sector," a spokesperson told *Nature*.

Davidovich disagrees. He says that, after accounting for inflation, research needs at least 10 billion reais to match the federal science spending levels of three years ago — so it is well under par.

To scientists' further dismay, the National Council for Scientific and Technological Development — a federal funding agency — has been further demoted within the new structure of the MCTIC. The council used to be an independent body whose president would report directly to the science minister; now it and other institutions such as the Brazilian Space Agency will report to a director who in turn will report to a vice-minister for science. "We [science] have been shoved to the fourth level," says Nader.

The political downgrading will make it harder for the council to fight for resources, say Nader and Davidovich. Last month, the Brazilian Society for the Advancement of Science and the Brazilian Academy of Sciences delivered a joint letter to Kassab protesting against the move. The minister said that he would review the changes in January.

For now, Brazilian science faces a cloud of uncertainty. As researchers have asked, the amendment could be altered to spare sensitive areas such as education and research. Senators might also try to spare a specific slice of the MCTIC budget that comes from the revenue of some businesses (such as oil and telecommunications), Davidovich thinks. It is hard to predict how the amendment might unfold in the Senate, he says. "But then again, someone has said that, in Brazil, even the past is unpredictable." ■

BIOLOGY

Geneticists seek keys to bat vocalizations

Bat 1K project will sequence DNA of more than 1,000 species.

BY RAMIN SKIBBA IN SAN DIEGO, CALIFORNIA

Some bats sing just as birds and humans do. But how they learn their calls and melodies is a mystery — one that scientists will try to solve by sequencing the genomes of more than 1,000 bat species.

The project, called Bat 1K, was announced on 14 November at the annual meeting of the Society for Neuroscience in San Diego, California. Its organizers also hope to learn more about the flying mammals' ability to navigate in the dark through echolocation; their strong immune systems that can shrug off Ebola; and their relatively long lifespans.

"The genomes of all these other species, like birds and mice, are well-understood," says Sonja Vernes, a neurogeneticist at the

Max Planck Institute for Psycholinguistics in Nijmegen, the Netherlands, and co-director of the project. "But we don't know anything about bat genes yet."

Some bats show babbling behaviour, including barks, chatter, screeches, whistles and trills, says Mirjam Knörnschild, a behavioural ecologist at the Free University of Berlin. Young bats learn the songs and other sounds from older male tutors. They use these sounds during courtship and mating, when they retrieve food and as they defend their territory against rivals.

Scientists have studied the vocal sounds of only about 50 bat species so far, Knörnschild says, and they know much less about bats' communication than birds'. Four species of bat have so far been found to learn vocal sounds from each other, their fathers and

other adult males, just as a child gradually learns how to speak from its parents (S. C. Vernes *Psychon. Bull. Rev.* <http://doi.org/btcf>; 2016).

SINGING LESSONS

Genetic studies have identified at least one gene in bats that is linked to speech and language, called *FOXP2* (M. Knörnschild *Curr. Opin. Neurobiol.* **28**, 80–85; 2014). The gene is also known to have a role in how people learn language, and in vocal learning in songbirds. Researchers working on the Bat 1K project expect to find that other genes are also involved in communication, and that many more bat species have the ability to learn songs, calls and other sounds. "It's not a rare trait," Knörnschild says. "I'm becoming convinced that there's a whole continuum in bat vocal learning."

Bats' echolocation ability has been studied for many years, partly because of its applications to sonar and radar. But scientists know very little about the acoustic communication and social behaviour that drive how bats learn their songs and sounds, says Michael Yartsev, a neurobiologist at University of California, Berkeley. The study of vocal learning in bats is "nearly completely untapped", he says — likening it to the state of research into birdsong 60 years ago. ■

The sparrow with four sexes

Ecologist Elaina Tuttle spent her life trying to understand the bizarre chromosome evolution of a common bird — until tragedy struck.

BY CARRIE ARNOLD

There was one sound that biologist Rusty Gonser always heard at Cranberry Lake — and there was one sound that he would never hear again.

Every summer for more than 25 years, Gonser and his wife, Elaina Tuttle, had made the trip to this field station in the Adirondack Mountains — a 45-minute boat ride from the nearest road. Now, as he moored his boat to the shaky wooden dock, he heard a familiar and short song that sounded like ‘oh-sweet-Canada’. The whistle was from a white-throated sparrow calling hopefully for a mate.

What he didn’t hear was the voice or laughter of his wife. For the first time, Gonser was at Cranberry Lake alone. Just a few weeks earlier, Tuttle had died of breast cancer.

Her entire career, and most of Gonser’s, had been devoted to understanding every aspect of the biology of the white-throated sparrow (*Zonotrichia albicollis*). Less than six months before she died this year at the age of 52, the couple and their team published a paper¹ that was the culmination of that work. It explained how a chance genetic mutation had put the species on an extraordinary evolutionary path.

The mutation had flipped a large section of chromosome 2, leaving it unable to pair up with a partner and exchange genetic information. The more than 1,100 genes in the inversion were inherited together as part of a massive ‘supergene’ and eventually drove the evolution of two different ‘morphs’ — subtypes of the bird that are coloured differently, behave differently and mate only with the opposite morph. Tuttle and Gonser’s leap was to show that this process is nearly identical to the early evolution of certain sex chromosomes, including the human X and Y. The researchers realized that they were effectively watching the bird evolve two sex chromosomes, on top of the two it already had.

“This bird acts like it has four sexes,” says Christopher Balakrishnan, an evolutionary biologist at East Carolina University in Greenville, North Carolina, who worked with Tuttle and Gonser. “One individual can only mate with one-quarter of the population. There are very few sexual systems with more than two sexes.”

The work helps to explain a long-standing puzzle for biologists. It shows how two identical chromosomes can evolve into distinct subtypes that can define the sexes of a species and their different behaviours. “These birds are an amazing system,” says Catherine Peichel, an evolutionary ecologist at the University of Berne. “The process of sex-chromosome evolution tends to erase much of the evidence of how it happened, so being able to watch the process in action is a huge benefit.”

What makes Tuttle and Gonser’s project even more unusual is the

accumulation of almost 30 years of data, “something that is almost unheard of in biology now”, says Melissa Wilson Sayres, a computational biologist at Arizona State University in Tempe. “Most people tend to jump from project to project.”

Gonser is determined to carry on the project. He returned to the field station this summer to continue Tuttle’s legacy and use this drab little garden bird to understand how sex chromosomes may have evolved. “Who knows — there might be many more species that have weird sex chromosomes and we’ve just never bothered to look,” he says.

A FLOCK OF IDEAS

Gonser’s mind moves at a mile a minute, his thoughts racing ahead of his grey Honda Civic. Dressed in basketball shorts and a T-shirt, and sporting days-old stubble, he speeds west on the I-70 interstate towards Indiana State University in Terre Haute, where he and Tuttle shared an office. As ideas tumble out — about new research projects, people to meet, restaurants to avoid — he pauses briefly and apologizes. “Sorry if I’m repeating myself. My brain hasn’t been working right since ...” he pauses briefly, and clears his throat, “since Elaina passed.”

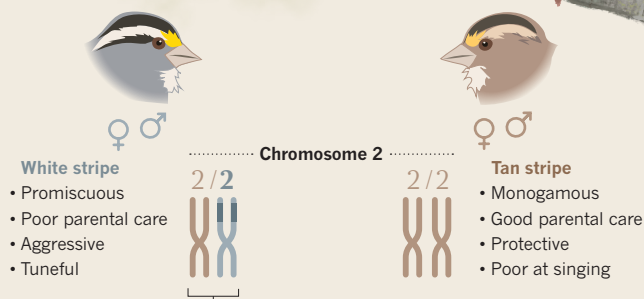
Tuttle and Gonser first met in 1991 at the State University of New York at Albany, where both were pursuing PhDs in ecology. Gonser was investigating the Puerto Rican frog *Eleutherodactylus coqui* and Tuttle was studying fish ecology at New York’s Finger Lakes. Long months in the field suited her love for the natural world, and it was there that she got to know the curious lifestyle of the white-throated sparrow.

The bird is relatively common in gardens across the eastern half of North America — and on first glance it is rather plain. All the birds have mostly tan and grey plumage, except for a patch of white under the chin and a bright pop of yellow between the eyes and beak. But closer scrutiny reveals the two types: some have white stripes on their heads, and some have tan stripes. And, as bird spotters and naturalists have long known, the two morphs behave in different ways.

The tan-striped birds are poor at singing, monogamous and fiercely protect their hatchlings from predators such as raccoons and snakes. The white-striped ones are aggressive, promiscuous, more cavalier about their offspring, and tuneful: Gonser says that they produce a more operatic refrain of oh-sweet-Canada. White-striped birds seem to mate only with tan-striped ones — a relatively unusual

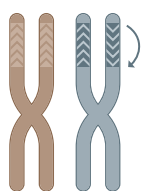
Opposites attract

The white-throated sparrow, common in eastern North America, has either white or tan stripes on its head. White-stripes mate only with tan-stripes, and vice versa. Why these two morphs behave so differently was a long-standing puzzle.



What's up with chromosome 2?

An ancient mutation inverted a large section of chromosome 2, creating a 'supergene'.



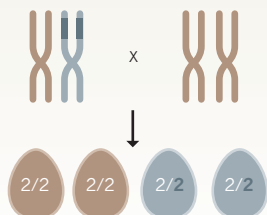
The inverted region is unable to exchange genetic information with its partner chromosome, so mutations accumulate.

The chromosomes diverged, driving different behaviours. This is thought to mirror the process by which mammalian (XY) and bird (WZ) sex chromosomes evolved.

Like having four sexes

Males always mate with females and

white-stripes almost always mate with tan-stripes (disassortative mating).



On average, half the offspring inherit the inverted version of chromosome 2, producing half tan-striped and half white-striped offspring.

phenomenon called disassortative mating (see 'Opposites attract'). Tuttle became interested; why do the two morphs behave in this way?

The literature already contained a huge clue. In 1966, ornithologist H. B. Thorneycroft published a paper in *Science*² pointing out the bird's unusual chromosome pair. Tan birds carry two identical copies of chromosome 2, but in white birds, one copy contains an inversion. It's as if someone had taken a pair of scissors, snipped out most of the chromosome and placed the DNA back in reverse. A chromosomal inversion this large was rare in vertebrates, Thorneycroft noted. It looked like disassortative mating maintained the two morphs in equal proportions in the population, because Mendelian inheritance ensures that, on average, half the offspring of a white-tan pair will inherit the inverted version of chromosome 2. But it would take more work to prove that this was true.

To Tuttle, it was a fascinating puzzle — a way to shed light on chromosome evolution, as well as the genes underlying social behaviour. In the early 1990s, however, it was too expensive and laborious to find answers by sequencing the bird's genome. So Tuttle initially focused on collecting more detail about their behaviour, such as how they selected mates and where they built nests. The goal was to understand what might affect offspring survival. She caught and tagged birds, drew blood samples and perfected the art of collecting semen. "Elaina was the best bird masturbator I ever met," Gonser says.

Gonser soon became drawn into the work. (The couple married in 1994.) After the birth of their son Caleb in 2000, the family began spending summers at Cranberry Lake, and they slowly learnt more about the sparrow. For a 2003 paper³, Tuttle used genetic analysis of nestlings to quantify the different reproductive strategies of the two morphs. She showed that nearly one-third of the offspring fathered by white-striped males were not born to females with whom they shared a nest. Tan males, by contrast, spent less of their energy finding extra partners and more time guarding their own, so they were more likely to have fathered the chicks they were raising. Although the two morphs went about reproduction in very different ways, both achieved equal reproductive success.

Six years later, the team secured a grant from the US National Institutes of Health to start genetic analyses. They mapped chromosome 2 in detail⁴ and found that the changes were not a single inversion as Thorneycroft had indicated, but actually a series of inversions within inversions that scrambled the order of genes. They identified several genes that seemed to be associated with feather colour and behaviour, and that might explain the differences between the two morphs. "It's really, really rare to find such a direct relationship between a set of genes and behaviours. That's what makes these birds so interesting to study," says Donna Maney, a neuroendocrinologist at Emory University in Atlanta, Georgia, who uses the white-throated sparrow as a model to understand how hormones affect the brain.

But then, in 2011, just as Tuttle and Gonser were ramping up their genome sequencing, a routine mammogram revealed a lump in one of Tuttle's breasts. A biopsy confirmed that it was invasive cancer. The couple was shocked — their son was only 11 — but a mastectomy and the drug tamoxifen seemed to contain the disease. Determined not to let the cancer control her life, Tuttle pressed on with her work.

DOUBLY DIVIDED

By this point, the researchers had come up with the concept that the birds were evolving a second set of sex chromosomes. "It's a bizarre idea, but it just makes sense from the data," said Alan Bergland, a geneticist who studies the molecular basis of evolutionary adaptation at the University of Virginia in Charlottesville. No other species was known to have one set of fully operational sex chromosomes and another pair that subdivided the species again on another aspect of mate choice.

Evolution has led to many weird and wonderful sexes and means of determining them. Some animals, such as reptiles, have two sexes and no sex chromosomes, whereas the freshwater protozoan *Tetrahymena thermophila* has seven sexes, each of which can mate with any type except its own. Two sexes with one set of sex chromosomes is the most



Elaina Tuttle dedicated her scientific career to studies of the white-throated sparrow.

common arrangement, and has evolved independently many times. But there's no reason a species can't have more sex chromosomes, Wilson Sayres says. "If you've got two genes linked together that can't cross over during meiosis, and one of them plays a role in sex determination, suddenly you can have a new sex chromosome."

Researchers believe that the X and Y chromosomes in many mammals, and the W and Z in birds, emerged from a major inversion on a normal pair of chromosomes that prevented them from swapping genetic material, or 'crossing over'. The suite of genes in the inversion was cemented together as a supergene that was inherited in one large chunk. On both the Y and Z chromosomes, the inversion shifted the location of a gene that determined male or female sexual development, respectively (male birds are ZZ and females are ZW). Over time, the Y and Z chromosomes accumulated mutations, because crossing over wasn't weeding them out. But, because all this happened in the distant evolutionary past, scientists had struggled to identify the precise steps involved.

The sparrows offered such a chance. The inversion on chromosome 2 doesn't include genes that determine sexual development — but it does contain some that affect the birds' reproductive behaviour. Over time, those genes diverged and drove the different characteristics of the two morphs. "Whatever genes control these behavioural differences will ultimately be traced back to this inversion," says Maney.

But to truly show that chromosome 2 was evolving like a sex chromosome, Tuttle and Gonser would need to demonstrate that the genes in the inversion were acquiring mutations much faster than those on other chromosomes. This would prove that the white and tan morphs were using true disassortative mating, so that all pairs were white-tan. Even a tiny percentage of white-white matings would allow the inverted chromosome 2 to undergo crossing over that would help to cleanse it of mutations, and undermine the idea that it was acting like a sex chromosome. Tuttle and Gonser would need to sequence the genomes of many birds, to show parentage and to compare mutation rates in the inversion with the rest of the genome.

Finding DNA to sequence was easy. By now, Tuttle had several freezers full of blood samples collected over the years from thousands of birds. But other biology was getting in the way.

In autumn 2013, a chronic cough sent Tuttle back to the doctor with what she thought was bronchitis. She learned that her

cancer had returned and spread to her lungs. Genome sequencing of her tumour revealed that her original cancer was a mosaic of cell types; the tamoxifen had tamed the hormone-sensitive cancer cells, but the insensitive ones had survived. She underwent more chemotherapy, which held the cancer in check.

INDIANA STATE UNIV.

THE LAST LAKE TRIP

By the summer of 2015, the team had gathered both the ecological and the genetic data that they needed and were putting the finishing touches to their big paper — when a routine scan revealed that Tuttle's tumours were growing again. Tuttle made the 13-hour drive from Terre Haute to Cranberry Lake between chemo treatments. "We knew she was sick, but we didn't realize just how sick she really was," says graduate student Lindsay Forrette. Soon, Tuttle was getting increasingly grim predictions about how much time she had left. "We were all thinking she was going to beat the cancer, and when we finally understood she wasn't ..."

Gonser's voice trails off. When Tuttle realized she would never return to Cranberry Lake, she broke down and cried.

In January 2016, when the paper was published in *Current Biology*, it showed unequivocally that chromosome 2 was evolving like a sex chromosome. White-white and tan-tan matings were exceedingly rare. Using the whole-genome sequences of 50 birds, the team demonstrated that the genes in the inversion were acquiring mutations much more quickly than elsewhere in the genome, a pattern that echoed the evolution of sex chromosomes in humans and birds.

In a press release, Tuttle said: "This is probably my pinnacle paper." It would also be her last. By spring, her health had worsened. Just five days before she died, Tuttle busied herself from her hospital bed, writing more papers and helping graduate students to analyse data. She died on 15 June. When Gonser announced the news, he received condolences from around the world. "Everyone who knew her, loved her. She was just that kind of person," Gonser says.

Tuttle leaves a rich research legacy that raises further questions, including whether this chromosomal system is ultimately destined to disappear. Balakrishnan says it is unsustainable. "That we never see systems with four sexes says that they're evolutionarily unstable and one of these alleles will ultimately go extinct." That's because in a four-sex system, each bird must work that much harder to

find a mate — a white female is not just looking for a male, she needs a tan one — so selection will favour a more advantageous two-sex system instead. Balakrishnan intends to use the white-throated sparrow to further tease apart the genetic and environmental factors that drove its evolution in the first place.

Gonser will keep going back to Cranberry Lake. He and his group want to better understand which genes control the sparrow's behaviour — from mate selection to parenting — and how those characteristics were affected by the inversion. They are using digital maps and satellite data to chart nesting sites, track tagged birds and build a richer set of behavioural data.

"There's a lot more information left in these birds," Gonser says. "And I think Elaina would like that we're trying to uncover their secrets." ■

Carrie Arnold is a writer based near Richmond, Virginia.

1. Tuttle, E. M. *et al. Curr. Biol.* **26**, 344–350 (2016).
2. Thorneycroft, H. B. *Science* **154**, 1571–1572 (1966).
3. Tuttle, E. M. *Behav. Ecol.* **14**, 425–432 (2003).
4. Romanov, M. N., Dodgson, J. B., Gonser, R. A. & Tuttle, E. M. *BMC Res. Notes* **4**, 211 (2011).

COMMENT

MATERIALS How printing could make thin photovoltaics mainstream **p.485**

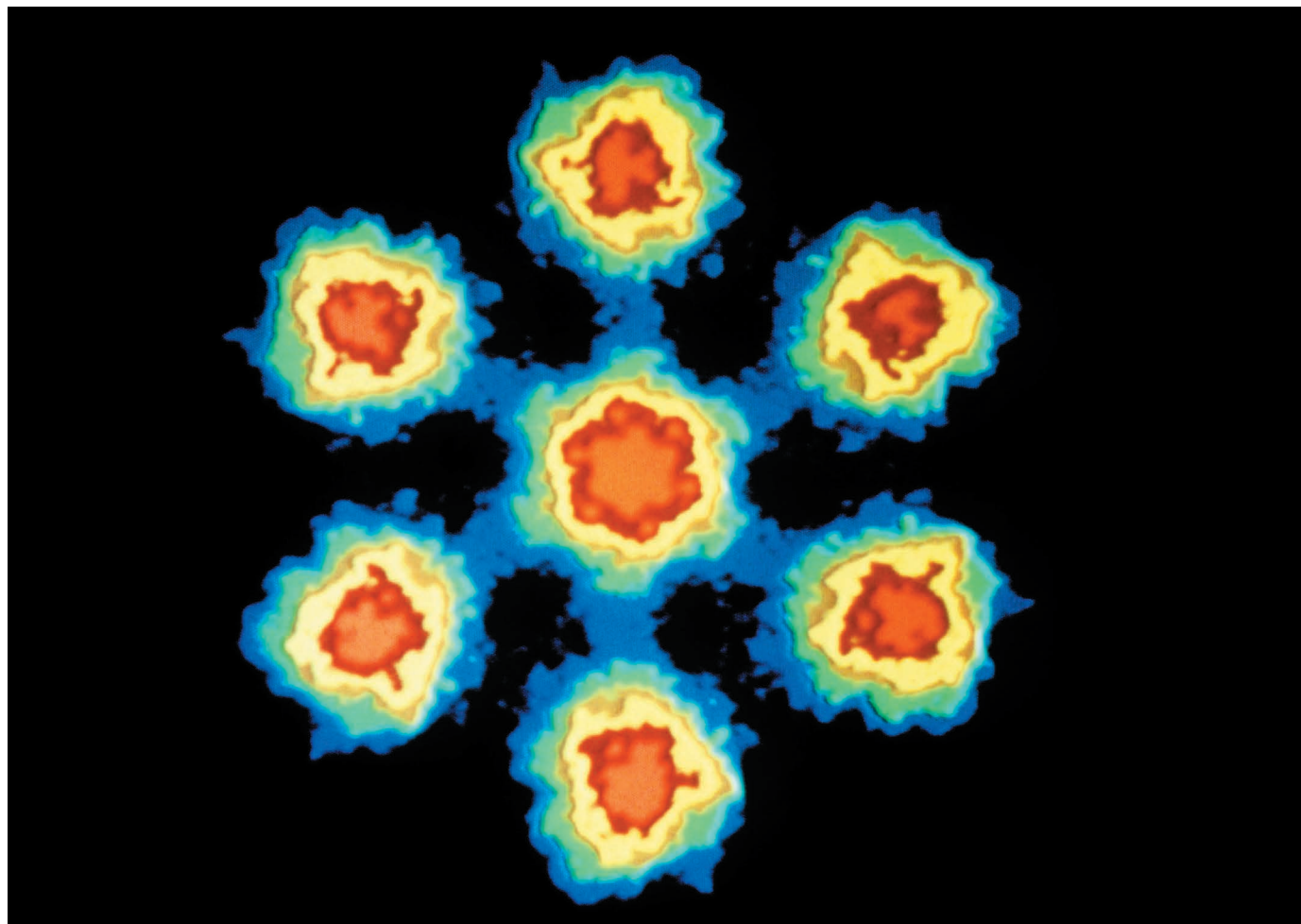
HISTORY Hymn to the pioneering Harvard women who catalogued the stars **p.491**

POLITICS A life on the front lines of battling pseudoscience **p.492**



GOVERNANCE Reflections on the long-range implications of Donald Trump's win **p.495**

MITSUO OHTSUKI/SPL



Scanning transmission electron microscopy has been used to reveal the hexagonal arrangement of the uranium atoms in a solution of uranyl acetate.

Fire up the atom forge

Rethink electron microscopy to build quantum materials from scratch, urge **Sergei V. Kalinin, Albina Borisevich and Stephen Jesse**.

Electron microscopy is on the brink of a transition. Soon the imaging tool could be used to create structures atom by atom. This sort of control over atomic architecture could transform our basic scientific understanding of materials and pave the way to new classes of devices for quantum computing, spin sensing and more.

Currently, the only way to build at the

atomic scale is with a scanning tunnelling microscope (STM). But this approach remains impractical and niche, 25 years after Don Eigler demonstrated it by spelling out the name of his company, IBM¹. An STM can move only surface atoms and structures that are stable only at low temperatures. The technique is time-consuming and it requires bespoke machines. It can now make

computing devices comprising several quantum bits (qubits), but not much more².

Electron microscopy is reliable and widely used to view thin sections of materials. It uses a beam of electrons to reveal the material's crystal structure (see 'The atomic forge'). However, the beam can nudge atoms out of alignment and alter the material's form. This is a problem for imaging, but could be a ►

THE ATOMIC FORGE

A **SCANNING TRANSMISSION ELECTRON MICROSCOPE (STEM)** fires a beam of electrons through a sample of material to pinpoint atoms and reveal the material's crystal structure. For imaging, the downside is that the beam can move atoms slightly. But if these modifications can be controlled, it could be a boon. New materials could be built atom by atom by controlling the electron beam precisely. Such bespoke materials might enable new classes of devices for quantum computing, spin sensing and more.

PRECISE BEAM

A beam delivers a prescribed dose of electrons to a particular site in a crystal for a certain length of time to move an atom as desired.

BEAM CONTROL

Open-source software gives researchers full control over the precision, velocity and power of the electron beam.

FEEDBACK SYSTEM

A powerful computer analyses the detector data in real time and decides how to control the beam to make the next atomic move.

FAST DETECTOR

Information about the sample's atomic structure is gathered in real time by detectors with many pixels.

NEW MATERIALS

By testing new atomic architectures, researchers will be able to design materials with specific properties and learn more about interactions between the beam and sample.

► boon for fabrication — as long as the modifications can be controlled.

For building at the atomic level, scanning transmission electron microscopy (STEM) seems to be the most promising form of the technique. An electron beam as fine as an atom is scanned through the specimen to probe its structure³. Several groups have used STEM to transform the compositions of layered materials in places^{4,5}, crystallize amorphous oxides⁶, move atoms between spaces in a lattice⁷ and knock out single atoms from one-layer materials.

Manipulating atoms with STEM will require three key developments, practical and theoretical, which we outline here. Practically, we need better real-time monitoring, feedback and beam control. These systems should use imaging methods⁸ based on diffraction signals. They will need detectors capable of being read out fast and ‘big data’ computation methods capable of handling high volumes and rates of data while performing analytics in real time. Materials scientists must also refine theories of how electrons and matter interact.

THREE WAYS FORWARD

Control the beam. In standard STEM imaging, an electron probe scans back and forth across the specimen. Commercial imaging systems typically trace only a zig-zag pattern over a rectangular area, because arbitrary beam paths can distort the image. To deliver a prescribed number of electrons to a specified location at a particular time demands full control over the position, velocity and power of the electron beam. Fine control is also useful where different parts of a sample react differently to the electrons or when only a portion of the field needs to be examined. Different scanning paths are useful, however. For example, spirals are faster to scan because they avoid abrupt changes in direction and deliver electrons more evenly⁹. The small, high-end manufacturer Nion in Kirland, Washington, has made its beam-control software open source so that researchers can develop models for different scanning strategies. Others should follow suit.

Monitor and feed back in real time. Irregularities and defects in lattices blur STEM images. To build atom by atom, irregularities need to be spotted and located quickly so that the fabricator can decide whether to correct, adjust or leave them in place. Currently, the best way to do this is to keep rescanning. This is too slow.

Researchers will need to be able to monitor and respond to detector signals in real time. Through a system of immediate feedback, the beam could be controlled to alter the rate and extent of modifications or reactions. Until that becomes available, regular rapid scans of small areas (such as 10-millisecond spiral

scans of 1 nanometre square) could be used. Images would need to be analysed in real time and algorithms developed for adjusting the scanning trajectories. The speed and quantities of data involved would require powerful computers and software.

In the long term, diffraction patterns could be used to track lattice structures, as is currently done in scanning-diffraction imaging. Collecting transmitted electrons on a pixelated detector instead of a single ring preserves more information. For example, the electron beam may scatter off new or missing atomic columns or be affected by symmetries and asymmetries in the surrounding crystal. Libraries of standard diffraction images can be consulted to show that a structure has been fabricated successfully.

Improve theory and modelling. Knowing more about how the electron beam interacts with matter is crucial for predicting how the fabrication system will behave. For instance, researchers need to know how slight changes in the positions of atoms in a column affect scattering. A better grasp is needed of coupled nonlinear effects, such as how the beam changes a material and thus how the two interact subsequently.

Greater understanding is needed of how electrons lose energy as they pass through the lattice (such as through collisions with atoms or by exciting acoustic modes). Such information will help to predict which strong interactions will knock out single atoms or displace them in cascades. The formation of new, stable configurations should be explored.

All this knowledge would allow fabricators to exploit subtleties such as the way a material responds to scan direction, path shape or the pulsing of the beam. The prize will be worth the effort: the researcher who can devise a scan path to steer an atom to a predefined location will be able to make materials for quantum computing.

NEXT STEPS

Realizing these opportunities will require interdisciplinary research. Those who work on electron microscopy, data analytics, image analysis, quantum computing, microfabrication, theory, molecular dynamics, computer visualization, artificial intelligence and instrument engineering must come together. Such diverse teams are rare, and each community has a different language and motivation.

To nucleate this effort, we propose that professional societies and think tanks host interdisciplinary workshops to discuss the feasibility and specific needs of atomic

SUMMARY

● Materials could be made from scratch by enhancing a microscope that uses an electron beam to image the structure of crystals.

● Making the electron beam programmable — to vary how many electrons are fired, to where and for how long — would allow atoms to be moved, added or knocked out.

● The structure of the material needs to be monitored in real time so that progress can be followed and mistakes corrected.

● Researchers need to know more about how electron beams interact with atoms in crystal lattices, including how energy dissipates.

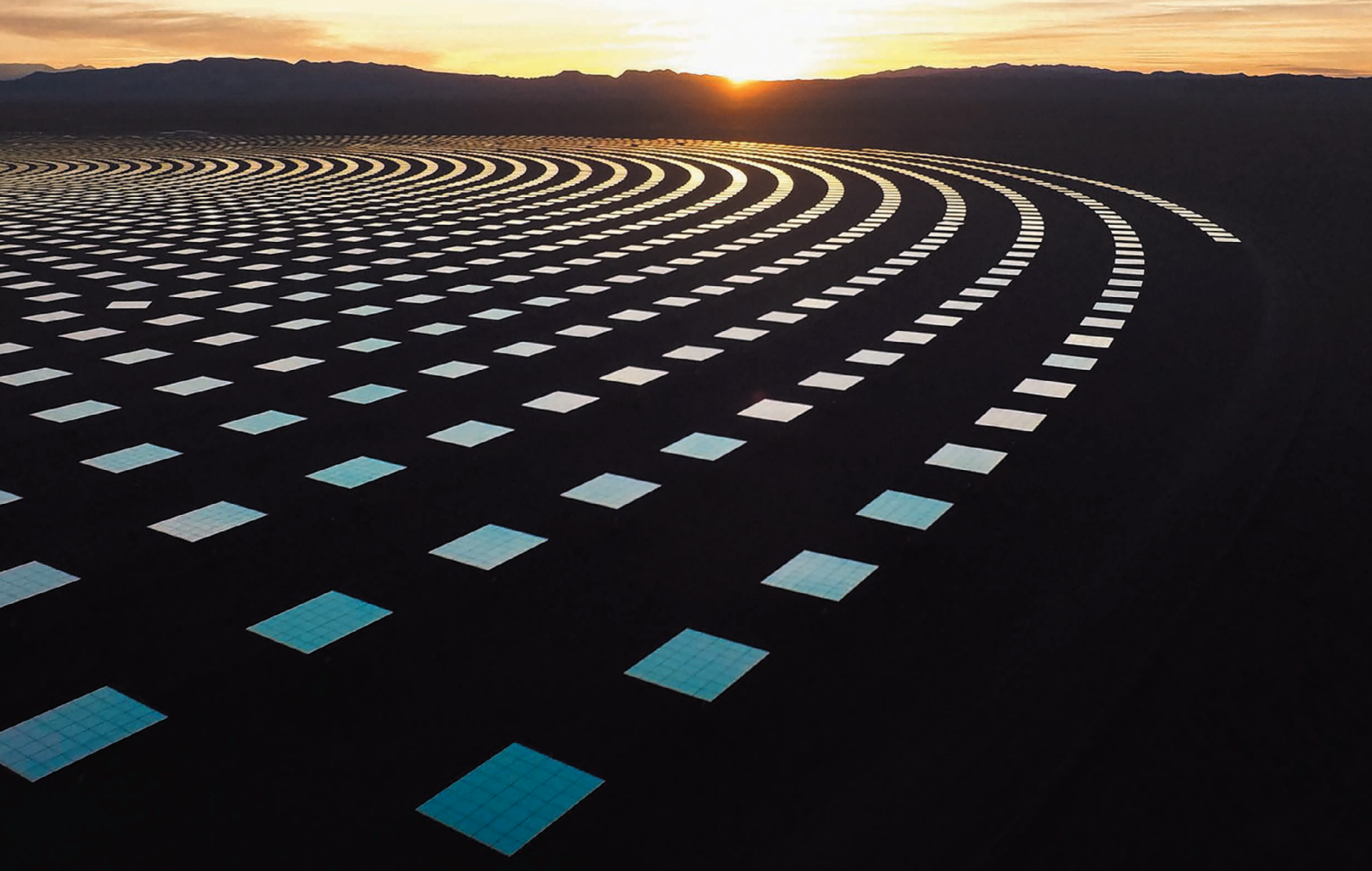
fabrication using STEM. Technical advances in the spatial resolution of STEM are equally important, laterally and along the beam.

Electron-microscopy centres should help the community share knowledge about observations, atomic configurations and beam-induced phenomena by setting up open and searchable databases with analysis tools. Common data formats and cloud computing facilities are needed to support data analytics. Our community should look to exemplars in other fields, such as X-ray, mass spectrometry and crowdsourcing platforms in high-energy physics, astronomy and genomics.

Let us make a start: the ability to build new forms of matter from the atom up will mark a new chapter of nanoscience. ■

Sergei V. Kalinin is distinguished research-and-development staff, and Albina Borisevich and Stephen Jesse are senior research-and-development staff, at the Institute for Functional Imaging of Materials, Center for Nanophase Materials Sciences, and Materials Sciences and Technology Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA. e-mail: sergei2@ornl.gov

1. Eigler, D. M. & Schweizer, E. K. *Nature* **344**, 524–526 (1990).
2. Fuechsle, M. *et al. Nature Nanotechnol.* **7**, 242–246 (2012).
3. Pennycook, S. J. & Nellist, P. D. *Scanning Transmission Electron Microscopy: Imaging and Analysis* 762 (Springer, 2011).
4. Lin, J. H. *et al. Nature Nanotechnol.* **9**, 436–442 (2014).
5. Huang, P. Y. *et al. Science* **342**, 224–227 (2013).
6. Jesse, S. *et al. Small* **11**, 5895–5900 (2015).
7. Ishikawa, R. *et al. Phys. Rev. Lett.* **113**, 155501 (2014).
8. Nellist, P. D. & Rodenburg, J. M. *Acta Crystallogr. A* **54**, 49–60 (1998).
9. Sang, X. *et al. Adv. Struct. Chem. Imaging* **2**, 6 (2016).



Conventional silicon solar cells account for more than 90% of global production, yet making them uses energy equivalent to about 10% of their lifetime output.

Print flexible solar cells

Pay more attention to developing thin, mass-produced, affordable photovoltaic devices, urge **Yi-Bing Cheng** and colleagues.

The future of solar energy depends on a union of new and old technologies. If photovoltaic (PV) devices that turn light into electricity could be mass produced with printing presses, as if they were newspapers or banknotes, they could be affordable and ubiquitous.

Conventional, silicon-based, solar panels are rigid and bulky. Small, thin and flexible PV devices on films are already being made that are lightweight and translucent. These use little material and can generate electricity in low light, even indoors. Integrating them into phones and watches, as well as walls and windows, would transform the world's energy generation, reduce pollution and mitigate climate change.

Yet flexible solar panels face several hurdles. Some are based on harmful substances such as heavy metals, and their manufacture uses hazardous solvents. Others are quick to degrade and inefficient at converting light into electricity. Printers used in the publishing, computing and

electronics industries struggle to print PV materials that need to be built with nanometre precision over many square metres. For all these reasons, printable solar cells are yet to find a foothold in electricity markets.

Most research and development investment goes into conventional silicon solar cells, which account for more than 90% of global production. Yet manufacturing these uses a lot of energy: equivalent to about 10% of the cell's lifetime output¹.

Printed solar cells won't become widespread until they are cheaper and safer to make. Researchers and businesses must work together to improve the efficiency, environmental impact and stability of these cells, scale up their manufacture and plan their market penetration.

PHOTOVOLTAIC PRIMER

Mass production at low cost is what the solar industry sorely needs. The power that a PV panel generates is proportional to the surface area exposed to sunlight. The world

consumes approximately 20,000 terawatt-hours of electricity each year². Meeting this need would require enough PV devices to cover around 100,000 square kilometres, an area about the size of Iceland.

Such production rates have yet to be achieved. Printed PV devices are typically made from many layers of material on a substrate of conductive glass or plastic. Each layer has a function: semiconductors or sensitizers absorb visible light, and other materials carry electrical charges to electrodes.

Many types of printed PV device are being developed. Some feature organic semiconductors such as polythiophenes. Others use light-absorbing dyes, including ruthenium-based polypyridines. And in quantum-dot solar cells, nanoparticles absorb light. Other examples feature semiconductors with a chalcogen element (sulfur, selenium or tellurium) or contain organic-inorganic light absorbers with a structure similar to that of the mineral

RUEBEN WU

perovskite. All of these are classed as thin-film solar cells.

The most efficient are the perovskite-based cells. The latest of these, with just a few years of research behind them (see 'Catching up'), convert 22% of incident solar power to electrical power³. This is more efficient than solar cells made from multicrystalline silicon. But perovskite cells cannot be rolled out commercially yet because they degrade under high humidity and heat.

FINE DETAIL

Printing layers that are nanometres to micrometres thick — uniformly and without pinholes, and over many square metres — is difficult. The electronics industry commonly uses screen-printing, feeding a paste through a perforated screen⁴. But the layers in printed circuit boards are hundreds of times thicker than those of PV devices. And turning materials into viscous pastes alters their physical and electrical properties.

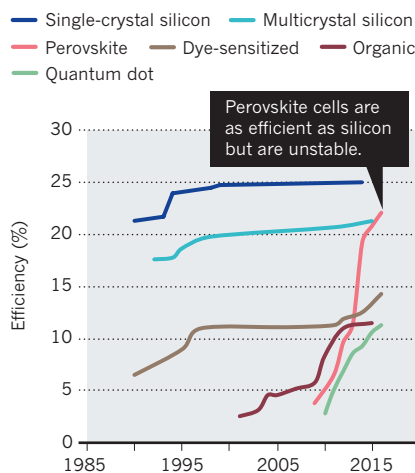
Other techniques for printing PV devices have been demonstrated in the lab over areas of about 10 square centimetres. These include feeding ink through a slit (slot-die printing), spray-coating the substrate, passing the substrate over a rotating cylinder (gravure printing) and moving a blade over the substrate through an ink supply.

Each technique has a downside. In slot-die printing, for example, the spacing of patterns is hard to control, and gaps reduce the active area of a panel. And in gravure printing, contact between the printer head and the substrate can damage underlying layers. Such drawbacks mean that printed solar cells are less than half as efficient as the best non-printed equivalents⁵.

To print thin, pinhole-free layers over more than one square metre will require intelligent, more-precise equipment and laser processing. An alternative approach would be to develop PV materials that work

CATCHING UP

Thin-film solar cells that can be processed as solutions and might be printed compete increasingly with silicon photovoltaics.



with existing industrial printing methods.

Printing requires that materials can be formed into a liquid, solution or paste. For PV devices, this means using either solutions of chemicals (polymers, dyes or hybrid perovskite, for instance) or dispersions of nanoparticles (such as quantum dots). But many of these can degrade over days to weeks if not properly sealed; and more-stable alternatives, such as silicon, are harder to print.

A balance must also be struck between the efficiency of a device and the environmental impacts of its manufacture. The most efficient thin-film solar cells include toxic or rare materials, such as cadmium, ruthenium and lead, as well as hazardous organic solvents. Indium, another rare element, is a common ingredient in transparent conductive films for PV devices, and its use is expected to rise.

Depleted mineral deposits and low rates of recycling threaten to exhaust indium stocks before the end of the century. Researchers

are looking for ways to use abundant materials that can be processed in solution to make efficient devices with little toxic waste. For example, carbon-based electrodes and layers use less precious metal⁶. At present, such designs are often less efficient than others.

WHAT NEXT?

The huge success of silicon panels has become a hurdle for emerging technologies. The manufacturers of silicon-based PV devices share materials, equipment and practices with sibling industries, such as computing. Developers of printable devices are isolated. And the maturity of the silicon industry means there is little urgency to develop alternatives. Capital investment and product commercialization are perceived as risky, given that printable PV devices are still in development.

Government funding is needed to catapult such devices from a nascent to a competitive state, as happened in China with the silicon PV industry. Market penetration should be developed in stages. Early printable PV devices should target weaknesses in silicon-based technologies, such as their poor performance in low light and their lack of portability. The next wave should complement silicon solar cells and, ideally, be integrated with them. For example, silicon-perovskite devices would harvest a greater fraction of incoming sunlight than silicon devices alone could do.

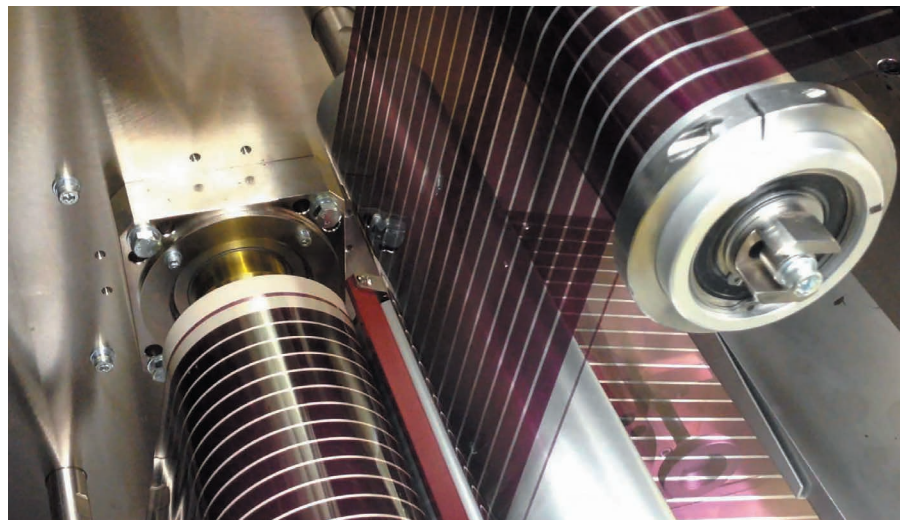
If printed technologies can capture 5% of the PV market, their advantages should ensure that they play an ever-increasing part in meeting growing demands for renewable energy. ■

Yi-Bing Cheng is professor of materials science and engineering at the State Key Lab of Advanced Technology for Materials Synthesis and Processing, Wuhan University of Technology, China, and at Monash University, Clayton, Victoria, Australia.

Alex Pascoe is a research fellow at Monash University, Clayton, Victoria, Australia.

Fuzhi Huang and Yong Peng are professors at the State Key Lab of Advanced Technology for Materials Synthesis and Processing, Wuhan University of Technology, China.

1. *Photovoltaics Report* (Fraunhofer Institute for Solar Energy Systems, 2016); available at go.nature.com/2eug7r
2. *Key World Energy Statistics* (IEA, 2016); available at go.nature.com/2ewux1w
3. *Best Research-Cell Efficiencies* (US National Renewable Energy Laboratory); available at go.nature.com/2exvaq0
4. Das, R. *Printing Equipment for Printed Electronics 2015–2025: Market Opportunities for Printing, Curing and Integration Equipment* (IDTechEx, 2015).
5. Ye, M., Hong, X., Zhang, F. & Liu, X. J. *Mater. Chem. A* **4**, 6755–6771 (2016).
6. Mei, A. et al. *Science* **345**, 295–298 (2014).



Roll-to-roll printing of organic photovoltaic cells



Some of the Harvard Observatory 'computers' in 1925. Annie Jump Cannon is seated fifth from left; Cecilia Payne is at the drafting table.

HISTORY

Women who read the stars

Sue Nelson delights in Dava Sobel's account of a rare band of human computers.

There are half a million photographic plates in the Harvard College Observatory collection, all unique. They date to the mid-1880s, and each can display the light from 50,000 stars. These fragments of the cosmos furthered our understanding of the Universe. They also reflect the dedication and intelligence of extraordinary women whose stories are more than astronomical history: they reveal lives of ambition, aspiration and brilliance. It takes a talented writer to interweave professional achievement with personal insight. By the time I finished *The Glass Universe*, Dava Sobel's wonderful, meticulous account, it had moved me to tears.

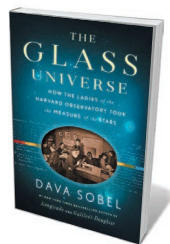
In the nineteenth century, US women had no vote (that would come in 1920). The careers open to the vast majority were in domestic service, farms, factories, schools or offices. Fortunately, Maria Mitchell, the first woman in the United States to discover a comet, left a guiding light. In 1865, she became the first professor of astronomy at Vassar College in Poughkeepsie, New York, a new institute of higher education for women.

Yet change was slow. When Edward Pickering became director of the Harvard Observatory in Cambridge, Massachusetts, in 1877, being a 'computer' was still predominantly a man's job. This low-paid but essential work involved noting a star's brightness, or

magnitude, on the photographic plates and computing its location in the sky. Pickering was especially interested in variable stars, whose light would brighten and fade over a specific period. These fluctuations, captured on the plates, required constant observation, but he couldn't afford extra staff.

Some years later, opportunity arrived in the form of the wealthy Anna Palmer Draper, widow of amateur astronomer and physician Henry Draper. She set up a memorial to fund his great ambition: to photograph the night sky and create a spectral catalogue of the stars. When starlight is split by a prism or spectroscope at the end of a telescope, it can produce a tiny grey smudge, millimetres long, on a photographic plate. The separated light, on closer examination, reveals vertical black spectral lines. This stellar barcode provides a star's chemical composition, colour and temperature.

By 1883, six women were already working



The Glass Universe: How the Ladies of the Harvard Observatory Took the Measure of the Stars

DAVA SOBEL
Viking: 2016

as computers at Harvard, a practice unique to the university. Within five years, the number of paid female computers, from a range of backgrounds, had risen to 14. Their efforts would be boosted by philanthropist Catherine Wolfe Bruce, who in 1889 donated US\$50,000 to the observatory, convinced that the introduction of photography and spectroscopy would advance the field.

The Glass Universe concentrates on a few of the Harvard computers. Williamina Fleming, a Scottish school teacher, arrived at the observatory in 1879, pregnant and abandoned by her husband. Pickering employed her as a maid and housekeeper before promoting her to computer in 1881. Fleming discovered more than 300 variable stars and, in 1899, became Harvard's first curator of astronomical photographs. She and Pickering revised the classification of stars using their spectral lines, labelling them alphabetically from the bluest, hottest stars to the coolest red stars.

Henrietta Swan Leavitt, who studied at another women's institution, Radcliffe College in Cambridge, began as a volunteer in 1895, gaining a permanent post in 1903. Five years later, she determined the relationship between luminosity and period of fluctuation for a class of pulsing variable stars called Cepheids. That enabled astronomers to calculate a star's distance from Earth, and ▶

COURTESY HARVARD UNIV. ARCHIVES

► in 1918 allowed Harlow Shapley, who would become Pickering's successor, to extend the boundaries of the Milky Way. Leavitt's work also laid the foundation for Edwin Hubble's finding, in 1925, that our Galaxy was not a lone 'island Universe' but one of billions. In 1929, Hubble used Leavitt's work and spectral shifts to show that this populous Universe is also expanding.

Annie Jump Cannon arrived in 1896, following studies in mathematics, physics and astronomy at another women's college, Wellesley, in the Massachusetts town of the same name. She refined and simplified the Pickering–Fleming system of stellar classification by proving that most stars were of a similar type. The International Astronomical Union adopted her system in 1922, three years after Pickering's death.

Cecilia Payne, educated at the University of Cambridge, UK, came to the observatory on a fellowship in 1923, under Shapley. As part of her PhD, she determined correctly that stars were predominantly composed of hydrogen and helium. In 1956, she became Harvard's first female professor of astronomy.

Payne provides some of the book's most touching passages. Sobel describes, for instance, Payne's distress at the untimely death of her friend and fellow computer Adelaide Ames in 1932, and her whirlwind romance with the astronomer and Russian émigré Sergei Gaposchkin, whom she married in 1934. Later that year, Payne won the inaugural Annie Jump Cannon Prize for contributions to astronomy, receiving \$50 and a gold pin, designed by Cannon, in the form of a spiral nebula.

These personal touches flesh out the women's lives. We are shown Payne, for instance, pawning her violin and jewellery to fund her research, and combining motherhood with astronomy. The low wages were a sticking point with many of them — Fleming often mentions it in her diaries. Yet, as Sobel shows, they clearly liked and admired Pickering and Shapley, who encouraged their work and facilitated their professional progress.

One of the last computers in the twentieth century, Radcliffe alumna Ellen Dorrit Hoffleit, gets a brief mention towards the end. She joined the observatory in 1928. In 2004, I met her at Yale University in New Haven, Connecticut, when I was making a radio documentary for the BBC on the Harvard computers. An energetic 96, Hoffleit was working on a paper about meteors. Like all of this band of remarkable women, she was unforgettable, as is this book. ■

Sue Nelson is an award-winning science journalist and broadcaster.
e-mail: sue@boffinmedia.co.uk

POLITICS

Life at the divide

Alison Abbott hails a memoir from Italian senator and biologist Elena Cattaneo, scourge of pseudoscience.

One hot day in August 2013, the phone rang in the laboratory of neuroscientist Elena Cattaneo at the University of Milan. The then-president of Italy, Giorgio Napolitano, wished to speak to her in Rome as soon as possible.

Cattaneo was already a minor national figure, known for her opposition to all forms of pseudoscience, however powerful the perpetrators. In one of her first public fights, in 2009, she criticized the Vatican for condemning the use of human embryonic stem cells in research. At the time of the call, she was in the middle of a battle to stop the government supporting a clinical trial of an unproven stem-cell 'therapy' run by the maverick Stamina Foundation in Brescia. The call made her wonder whether she had overstepped some mark.

Napolitano had something quite different to discuss at his residence in Rome, the baroque Quirinal Palace. He told Cattaneo, then aged just 51, that he wanted her to become a senator for life: what did she think?

"It was as if my every atom had been paralysed by the question," she recalls in her book *Ogni giorno: Tra scienza e politica (Every Day: Between Science and Politics)*. "Seconds passed without me being able to line up a meaningful sentence. The President realised, and touching my arm as if to shake me, he jokingly said: 'Professor, can I get you a cordial?'"

Cattaneo works on Huntington's disease, an incurable genetic disorder in which brain

Ogni giorno: Tra scienza e politica (Every Day: Between Science and Politics)

ELENA CATTANEO
Mondadori: 2016.

cells progressively die. Stem-cell research has provided new insights into the disease, which is why she had fought so hard for the right to do it responsibly.

Those battles stole precious time and energy. Would being a senator for life — with the inevitable widening of her repertoire of fights — drain too much from her scientific life?

Ogni giorno is a fascinating account of how a scientist entered the messy business of politics (arguably more so in Italy than in other rich countries) and learnt to survive. Cattaneo accepted the position. As one of only a few appointed by the president, the role signifies a person's importance to national culture. But she set boundaries. She would not give up her lab, and in the Senate she would engage mainly with themes relating to science. She would apply the same rigour to evidence supporting a political hypothesis as she would to that supporting a scientific one. She would not vote if such evidence were missing, and she would vote only according to her conscience, never along any party lines.

In August that year, she was sworn in along with conductor Claudio Abbado (who died five months later), architect Renzo Piano and Nobel-prizewinning physicist Carlo Rubbia. She threw herself into action, appointing a small team to help with research and administration, and organized a series of workshops with invited scientists from around the world. The aim was to educate parliamentarians in themes such as science, health and innovation, neuroscience and cell therapy.

She quickly learnt about Senate investigations and organized one into the Stamina debacle, a process that took more than a year. Published in February 2015, the report assigns responsibility for every step in the affair that allowed public hospitals to host the treatment, the government to support it and numerous courts to rule that patients had the right to it — even though it had been condemned as dangerous by the Italian health authorities.

Her repertoire of battles, as anticipated, did expand and she found herself on steep learning curves. One of these was genetically modified (GM) crops. In 2015, Italy's parliament debated the sensitive issue of whether each European Union member state should have the right to ban cultivation in their territories



Elena Cattaneo.

ANTONIO SCATTOLINO/CONTRASTO/EVINE

of GM crops that had been approved by EU authorities. Giving evidence at a Senate hearing, veterinary scientist Federico Infascelli described experiments in which he had apparently showed harmful effects in animals that were fed GM crops.

Cattaneo spent the weekend poring over his papers. She could not understand the data, and discussions with Infascelli did not help. So she prompted investigations into his work, which eventually saw some of the papers retracted. Her team went on to produce a 1,556-page document for the Senate compiling nearly all available scientific evidence about GM crop safety. For her pains, Cattaneo was demonized by some in the Senate as a lobbyist for the agricultural biotechnology firm Monsanto.

She has confronted many other issues in the Senate — for example, the continuing need for animals in research, the non-transparency of some government decisions relating to science, the murky proposal for a major research centre to be built on the site of the 2015 Expo in Milan, and issues such as research financing.

Ogni giorno is a dispatch from a strong woman fighting in a hostile environment. She sees an anti-intellectual tendency in her parliament, an anti-scientific attitude in the Italian judiciary and a reluctance of most scientists in the country to speak up in scientific controversies — “leaving the floor open to pseudoscientists and charlatans”.

Inevitably, she has made enemies. But she has made friends, too. In June, a total stranger, an accountant from the small provincial town of Molinella near Bologna, left his entire fortune of more than €1.5 million (US\$1.6 million) to her to distribute to research in ways she sees fit.

Having reported these various stories for years, I was fascinated to learn from the book how many more weapons one has when fighting battles for science from within a political system — but just how taxing those battles can be.

Cattaneo might turn out to be the last senator for life to be appointed. There will be a referendum on the Italian constitution on 4 December. A ‘yes’ vote will lead to a smaller, less powerful Senate in which the president will have the right to directly appoint a handful of senators for only a seven-year period. ■

Alison Abbott is Nature's Senior European Correspondent.



A Glandville fritillary butterfly (*Melitaea cinxia*), subject of a unique long-term data-collection effort

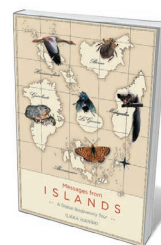
ECOLOGY

Winged insights

H. Charles J. Godfray is inspired by the scientific memoir of late island ecologist Ilkka Hanski.

Combining a personal memoir with serious discussion of a scientific subject is a difficult literary trick. The Finnish biologist Ilkka Hanski succeeded with aplomb in his last book, *Messages from Islands*, in which each chapter begins with insights from an island that moulded his thinking about ecology, evolution and conservation. Hanski, one of the foremost ecologists of his generation, died in May (A.-L. Laine *Nature* 534, 180; 2016).

Finland is a land of lakes and islands, so perhaps it is not surprising that Finnish ecologists are drawn to investigating how populations and communities persist in fragmented habitats. Hanski is most celebrated for developing the ecological concept of a metapopulation — a population of populations connected by dispersal — and its applications to conservation.



Messages from Islands: A Global Biodiversity Tour
ILKKA HANSKI
University of Chicago Press: 2016.

There are several types, but a classical metapopulation is sometimes likened to a collection of “blinking lights”, with individual short-lived populations winking in and out of existence while the whole ensemble persists.

Hanski explored the concept through his 25-year, and ongoing, study of the Glandville fritillary (*Melitaea cinxia*) in the Åland Archipelago between Finland and Sweden. This checkerspot butterfly has exacting habitat requirements: it occupies a fluctuating

▶ number of the small woodland meadows that constitute a habitat archipelago within the geographical archipelago. The meadows are so tiny that they support only a small butterfly population; each has a high risk of extinction every year.

Hanski, his colleagues and an ever-changing army of students surveyed all 4,000 or so meadows, which support 400–800 populations each year. Through this and many experiments, such as quantifying rates of dispersal between patches, they constructed a model of the butterfly's metapopulation — the most detailed and satisfying description of such a population structure currently available, by some distance.

Åland, of course, is the basis of a chapter. We learn how Hanski drew on the deep knowledge of Finnish butterflies and moths that he gained as a keen teenage naturalist in his search for a suitable model system. As elsewhere in the book, he explores much broader questions in population biology — for example, how agriculture, deforestation and other types of habitat change have shattered what would once have been contiguous habitats. Many more species live in metapopulations today than in the past, so understanding them can help in the design of protected areas, increasing the chance that a threatened species will persist.

Hanski's passionate concern for biodiversity is evident throughout the book. The chapter on Madagascar discusses the island's community of nearly 300 species of dung beetle (the subject of Hanski's doctorate and an abiding passion throughout his life). Madagascar has experienced extensive deforestation in the past century, making it likely that some of the dung-beetle species known from nineteenth-century collections are now extinct. More positively, Hanski and his students have documented how some species have



Biologist Ilkka Hanski, who died this year.



A few of the 300 dung-beetle species endemic to Madagascar.

evolved to utilize the dung of introduced livestock such as cattle.

This leads to a more general discussion of how biodiversity is generated, for example drawing lessons from Darwin's finches and pollinating bees about when natural selection favours specialization or being a jack-of-all-trades. That, in turn, leads back to the Glanville fritillary, in which Hanski has demonstrated genetic differences in the colonization ability of butterflies from the core and periphery of the metapopulation.

Haminanluoto is the smallest island discussed in the book. This 2-hectare islet is on the south coast of Finland near the village of Hanski, named after a seventeenth-century ancestor of the ecologist. At the age of 14, Hanski counted 15 species of bird on the island. Repeat counts in 2003 and 2013 found about the same number of species, although some of the original species had gone and new ones had colonized the island. Hanski muses on how human activities have affected the natural dynamics of species turnover. This leads to thinking about

the future potential impacts of climate change.

Here again, Hanski's own experiences are illuminating.

As a doctoral student he took part in an expedition to the virgin forests of Gunung Mulu in Sarawak, on the north coast of Borneo. He studied the dung-beetle communities living at different altitudes and found that species composition changed as you moved up the mountain. Thirty-five years later, he returned and repeated the exercise. He found the same sequence of species along the transect, but the whole distribution had shifted upwards, probably owing to global warming. Hanski laments another difference: Gunung Mulu is itself now an island in a sea of oil palm, no longer part of a much more extensive forest.

Hanski wrote *Messages from Islands* after being diagnosed with cancer. His book does not shy away from the challenges that face humanity or the threats to biodiversity. Yet it is an uplifting read, revealing the author's humanity and deep love of the natural world. I cannot think of a better book to give to a general reader who wants an insight into modern ecology and how ecologists go about their trade. ■

H. Charles J. Godfray is in the Department of Zoology at the University of Oxford, UK.
e-mail: charles.godfray@zoo.ox.ac.uk

➔ NATURE.COM

For more on science in culture see: nature.com/booksandarts

Correspondence

Trump: China could take lead on climate

Donald Trump has questioned the science behind climate change and committed the United States to withdraw from the Paris agreement, raising concerns that the US contribution to international climate funds will cease (see also *Nature* <http://doi.org/bs57>; 2016). This could cause some of the least-developed countries to lose interest in the Paris process and make many of its essential elements much harder to implement.

The Paris agreement is supposed to start its first review in 2018 — a process that is already fraught with political difficulties that call for strong leadership from the United States, Europe and China. If the United States pulls out and Europe is distracted by problems such as Britain's upcoming exit from the European Union, it could cement China's position as a leader and shaper of global order.

David Victor *University of California, San Diego, California, USA.*
david.victor@ucsd.edu

Trump: keep climate plans to boost jobs

The economic insecurity of US voters contributed to the election of Donald Trump as president. Alarmingly, he has pledged to support coal usage, withdraw the US commitment to the 2015 Paris climate agreement and reverse the Clean Power Plan, which aims to regulate carbon dioxide emissions from power plants. Although the scientific community cannot back away from climate change, researchers could pay more attention to the economic needs of voters and develop strategies that manage climate risks while bolstering employment and the US economy.

For example, cheap domestic natural gas delivered by fracking has created economic benefit and made the United States an energy

exporter. As long as operators control fugitive emissions, gas from fracking will drastically reduce US carbon emissions. If appropriate environmental and social control were to be combined with carbon capture and storage — a technology best implemented by repurposed oil companies — cheap gas could provide both an economic and a climate solution.

The few countries that have already decarbonized their electricity systems have done so by relying on nuclear power. A revitalized US nuclear industry could deliver large amounts of safer and cheaper carbon-free nuclear power — and provide jobs in design, construction and maintenance.

A focus on emissions and technology coupled to economic development, not to ideology and causation, might be more successful with future voters.

Jane C. S. Long *Oakland, California, USA.*
janeclong@gmail.com

Trump: unlikely to deter clean energy

It is unlikely that the US election result alone can derail climate-change mitigation and clean-energy solutions (see also *Nature* **539**, 337–339; 2016). The global momentum in policy and technology is in large part beyond the reach of the next US administration.

Policies to support emissions reduction and clean-energy expansion are already in place around the world. As well as their impact on climate change, these cut pollution, create jobs and boost national competitiveness. These are strong reasons against policy reversal in China and elsewhere.

Some US federal regulations, such as fuel-economy standards, are not easily overturned. In the electricity sector, several US states are well positioned to make up for changes to federal policies that could otherwise limit

climate progress, at least over the next 4–8 years.

This policy momentum is partly due to the success of past policies in driving down the cost of renewables, through technological innovation and economies of scale in private firms. This has encouraged further growth in markets and in worldwide commitments to support decarbonization, which can continue despite new uncertainties about US federal policy.

Jessika Trancik *Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.*
trancik@mit.edu

Trump: threats to space science?

President-elect Donald Trump has declared his support for NASA's exploration mission and its human space-flight programme in particular. As a professor of planetary and space sciences, I am concerned that his opinion may be coloured by the flag-waving opportunities that these present, and that he could be interested in territorial rather than global benefit (see also *Nature* <http://doi.org/bs58>; 2016).

I am also concerned about statements he has made regarding public–private partnerships, and increasing the role of the private sector in space exploration. As a businessman, he may decide to sell off some of NASA's assets to the private sector, without sufficient thought for safety and governance.

In my view, the international collaborative efforts that are currently in place are the correct way to progress. More-stringent visa regulations could make it harder for UK researchers to take up fellowships and research positions in the United States, hampering our partnerships in space missions and restricting the sharing of data and technology.

Monica Grady *Open University, Milton Keynes, UK.*
monica.grady@open.ac.uk

Illness should not curtail PhD funding

As PhD students who were diagnosed with cancer during our training, we have first-hand experience of the financial consequences of taking time out for treatment and recovery. Cutting a student's stipend because of serious illness is standard practice in the United Kingdom, compounding the combined stresses of delays to the research project and of the condition itself.

In 3–4 years, UK doctoral students are expected to master new techniques and to generate a solid body of results for research papers and a thesis. Their university and funding contracts reflect this 'apprentice' status and come with restricted rights.

Training grants from the UK research councils typically allow 13 weeks of paid sick leave; the studentship is suspended if an illness is expected to last longer (see www.rcuk.ac.uk/funding/grantstcs). Other funders may not cover sick leave at all, and most universities do not step in to help.

As a result, a student who is ill can be forced to continue working, to rely on contributions from friends and relatives, or to quit their PhD — all unacceptable 'punishments' in our view. We suggest that training-grant contracts should be more similar to academic staff contracts in providing at least six months of paid sick leave.

Edwin S. Dalmaijer *University of Oxford, UK.*

E. A. Claudia Pama *University of Cambridge, UK.*

Stella Prins *University College London, UK.*
edwin.dalmaijer@psy.ox.ac.uk

CORRECTION

Lihua Yang's Correspondence (*Nature* **539**, 168; 2016) should have referred to possible rewards to scientists publishing in 'top' international journals, not just in Chinese journals.

Ephrin Bs and canonical Reelin signalling

ARISING FROM A. Sentürk, S. Pfennig, A. Weiss, K. Burk, & A. Acker-Palmer *Nature* **472**, 356–360 (2011); doi:10.1038/nature09874

During cortical development, cell–cell interactions between ephrin receptors and ephrin B ligands are important for neuronal migration, axonal and dendritic growth, synaptogenesis, and pruning¹. Deficiency in Reelin (Reln) disrupts radial pyramidal neuron migration and results in cortical layer malformation¹. Recently, Sentürk *et al.* claimed that EphB receptor–ephrin B ligand (EphB–Efnb) and Reln pathways interact genetically and that layering defects in triple *Efnb1;2;3*^{−/−} mutants and *Efnb3*^{−/−}; *Reln*^{+/-} compound mice are similar to those present in *reeler* mice². Here we demonstrate that, contrary to this previous report, neither the compound mutant *Efnb3*^{−/−}; *Reln*^{+/-} nor the triple-mutant *Efnb1;2;3*^{−/−} show defects of cortical layering or elevated Dab1 levels. Although protein–protein interactions between Efnb ligands and Reln have been confirmed, they do not affect the radial positioning of migrating neocortical pyramidal neurons. There is a Reply to this Comment by Sentürk, A. *et al.* *Nature* **539**, <http://dx.doi.org/10.1038/nature/20130> (2016).

The similarity of the *reeler*-like phenotype in mice that lack the Reln receptors Vldlr and Apoer2, the intracellular kinases Src and Fyn, or the adaptor protein Dab1 led to the identification of the canonical Reln pathway: the binding of Reln to its receptors induces Dab1 phosphorylation (pDab1) via Src-family kinases, regulating actin and tubulin dynamics of the cytoskeleton necessary for orderly radial migration of cortical neurons. Degradation of pDab1 occurs through ubiquitination, so levels of unphosphorylated Dab1 in *reeler* mice are elevated. Dab1 binds to the intracellular NPXY motif of Vldlr and Apoer2 via its phosphotyrosine binding domain, yet it is unclear how Src-family kinases are recruited to phosphorylate Dab1. Sentürk *et al.* suggested that this role was filled by ephrins². Ephrins are membrane-anchored proteins that control axon guidance and other forms of cell–cell communication, including at the synapse. Cellular communication is bidirectional through kinase activation in the Eph-expressing cell and effector recruitment on the Efn-expressing cell³. Sentürk *et al.* reported that the ectodomains of Efnb2 and Efnb3 bind to Reln². We confirmed that Efnb3 binds to Reln, whereas Efnb1 does not⁴. In addition, genetic interaction of *Efnb* genes with *Reln* was implicated through the analysis of compound *Efnb3*^{−/−}; *Reln*^{+/-} animals, which Sentürk *et al.* reported to exhibit a similar phenotype to the *reeler* mouse². On the basis of their results they concluded that recruitment of Src or Fyn requires Efnb ligands, making Efnb1–3 essential for canonical Reln signalling and neuronal migration during brain development². We have been unable to reproduce key elements of the mechanism reported by Sentürk *et al.*²

We previously reported that Reln also directly binds to EphB receptors⁴. Furthermore, Reln induces EphB2 clustering, tyrosine phosphorylation and degradation, though this is less efficient than by Efnb1. This signalling mechanism was independent of the canonical Reln receptors Apoer2 and Vldlr or of Dab1 phosphorylation⁴. To determine whether EphB receptors play a similar role in the radial migration of cortical pyramidal neurons as described for Efnb ligands (particularly for *Efnb3*^{−/−}; *Reln*^{+/-})², we examined the brains of *Ephb2*^{−/−}; *Reln*^{+/-} and *Efnb3*^{−/−}; *Reln*^{+/-} mice (*Efnb3*^{−/−} mice have been described in detail before⁵). In contrast to Sentürk *et al.*², we were unable to detect any malformations of cortical pyramidal neuron layering in either of these genotypes. Nissl staining (Fig. 1a–d), NeuN (Fig. 2a), Brn1 (Fig. 1i–l, q–v), and Tbr1 (Fig. 1m–p) immunoreactivity and analysis of the distribution and polarity of Thy1–GFP–M-positive neurons (Fig. 1e–h and Supplementary Fig. 1) failed to reveal any obvious cortical layer malformation. In addition we did not find the described

difference in the number of NeuN-positive neurons in the cortical marginal zone (Fig. 2a, c), the hippocampal CA1 stratum oriens (Fig. 2b, c) or the cerebellar molecular layer² (Fig. 2b, c) when comparing *Efnb3*^{−/−}; *Reln*^{+/-} mice to their *Efnb3*^{−/−}; *Reln*^{+/+} littermates.

During development, Cajal–Retzius cells secrete Reln and express a set of different EphB and Efnb proteins on their surface. Early in cortical development, Cajal–Retzius cells in *Ephb1;2;3*^{−/−} mutant mice disperse less efficiently, resulting in large gaps between neighbouring Cajal–Retzius cells in the marginal zone, most prominently in the caudal region of the hemispheres⁶. We therefore analysed the layering of the caudal cortex of *Efnb3*^{−/−}; *Reln*^{+/-} mice, but could not find any differences in this anatomical locale either (Fig. 1i, k, m, o, q–s). However, we recorded a non-Mendelian inheritance pattern in *Efnb3*^{−/−}; *Reln*^{+/-} × *Efnb3*^{−/−}; *Reln*^{+/-} and *Ephb2*^{−/−}; *Reln*^{+/-} × *Ephb2*^{−/−}; *Reln*^{+/-} genotype crosses at weaning age, with only 8% (12 of 150, $P < 1.5 \times 10^{-6}$ by χ^2 test) or 16% (27 of 174, $P < 0.005$, χ^2 test) of double-knockout mice surviving, respectively. The survival of mice homozygous for the *reeler* mutation in the *Reln* compound *Efnb3*^{−/−} line was significantly lower than in the *Ephb2*^{−/−} line ($P < 0.05$, two-tailed Fisher's exact test).

We next attempted to reproduce the reported *reeler*-like phenotype of the *Efnb1;2;3*^{−/−} triple knockouts². To generate *Efnb1;2;3*^{−/−} triple knockouts we used a *Nestin-cre* transgenic line, as was used by Sentürk *et al.*², and combined it with *Efnb1*^{lox/y} and *Efnb2*^{lox/lox} conditional alleles and an *Efnb3* null allele. In the developing central nervous system, the *Nestin* promoter is active from embryonic day (E)7.75 onwards⁷ and *Nestin-cre*-driven recombination can be detected as early as E11 (ref. 8). Our analysis of *Efnb1;2;3*^{−/−} triple knockouts did not reveal abnormal cortical pyramidal neuron migration (Figs 1s, v, 2a) or canonical Apoer2/Vldlr-dependent Reln pathway disruption (Fig. 2d, e).

Analysis of *Ephb1;2;3*^{−/−} triple knockouts for the three predominant EphB receptors expressed in the brain also failed to reveal cortical layering malformations (Figs 1s, v, 2a) or canonical Reln pathway disruption (Fig. 2d, e).

During the phase of pyramidal neuron migration (E11–E17) Reln is produced by Cajal–Retzius cells at the pial surface, whereas Apoer2 and Vldlr are expressed by migrating pyramidal neurons. Temporal expression of *Ephb* and *Efnb* mRNAs in the neocortex, as revealed by the Allen Developing Mouse Brain atlas⁹, showed expression of *Efnb1*, *Efnb2*, *Ephb1* and *Ephb2* from E11.5–E18.5. *Efnb3* and *Ephb3* expression was extremely low in the cortex. By contrast, almost all *Ephb* and *Efnb* genes were highly expressed in the developing hippocampus, where they are essential for proper layering and axonal and dendritic outgrowth^{3,10,11}.

The *Efnb3*^{−/−} mutant mouse strains used by Sentürk *et al.*² and our study were generated by different knockout strategies. Sentürk *et al.*² used the *Efnb3*^{−/−} mutant described previously¹², in which a *lacZ* knock-in at the initiating methionine codon in the first exon of *Efnb3* generates a true null allele. By contrast, the *Efnb3* mutation used in this study involves the insertion of a *neo* cassette into *Efnb3* intron 4 (*Efnb3*^{neo/neo})⁵, which causes the loss of any detectable protein expression¹³. Phenotypic expression of the gene defect in both strains is indistinguishable: examples include a unique locomotor defect, cortico-spinal projection defect and impaired long-term potentiation^{5,12,14,15}. Taken together, the data are consistent with the *Efnb3* mutant developed previously⁵ being a null allele, rather than merely hypomorphic. There is ample evidence that differences in genetic background can contribute to the severity of mutant

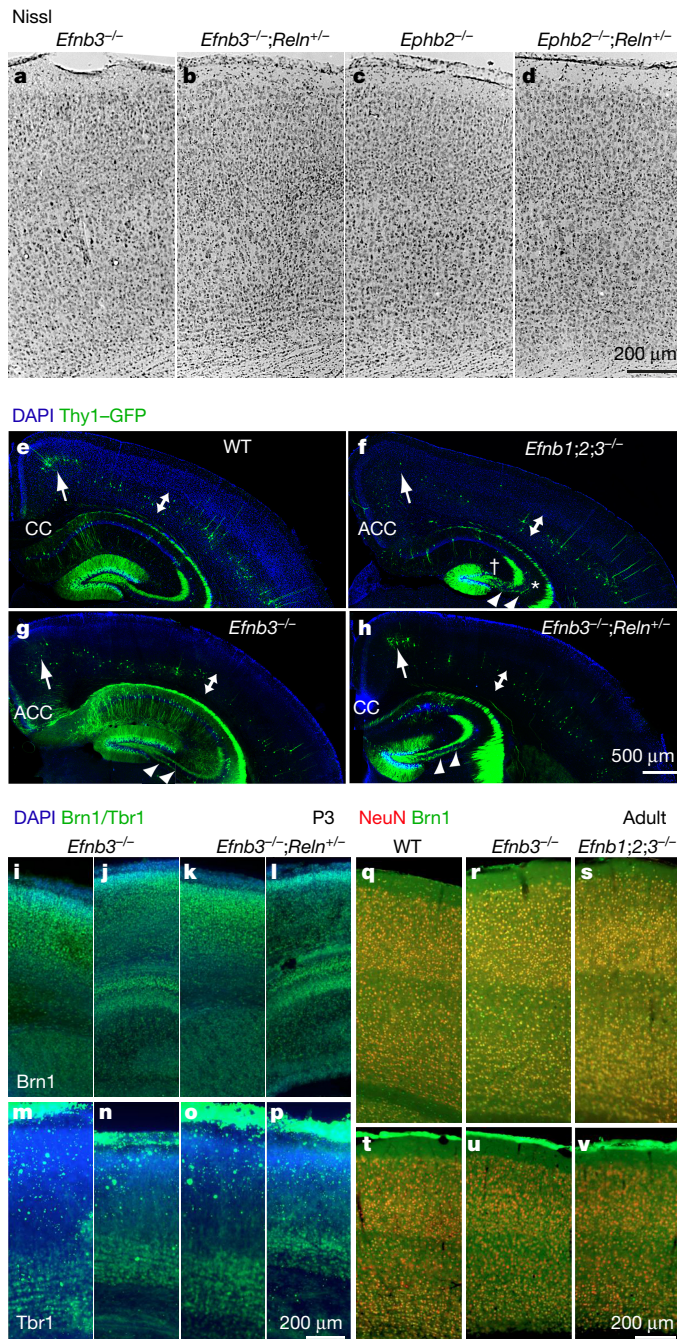


Figure 1 | *Efnb* ligands are not essential for *Reln*-guided radial pyramidal neuron migration. **a–d**, Nissl staining in the somatosensory cortex of adult mice with the indicated genetic mutations. **e–h**, Thy1–GFP-labelled layer V neurons of adult mice (double-headed arrows). In all genotypes a few ectopic cells were found in the retrosplenial area (arrows). Previously described prolonged fibres in the hippocampal infrapyramidal bundle (arrowheads)¹⁰, dispersion of CA3 pyramidal cells (asterisk), reduced volume of the granule layer (dagger)^{4,11} and agenesis of the corpus callosum (ACC, found occasionally) are indicated. **i–p**, Brn1 (**i–l**) and Tbr1 immunoreactivity (**m–p**) of P3 cortices. **q–v**, Brn1 and NeuN in the adult cortex. CC, corpus callosum. Images are caudal (**i**, **k**, **m**, **o**, **q–s**) or rostral (**j**, **l**, **n**, **p**, **t–v**) of the somatosensory cortex.

phenotypes. However, the effects of *Reln* knockout on neuronal layering and the effects of *Efnb3* knockout on locomotion and mossy fibre pruning are very robust in a variety of genetic backgrounds.

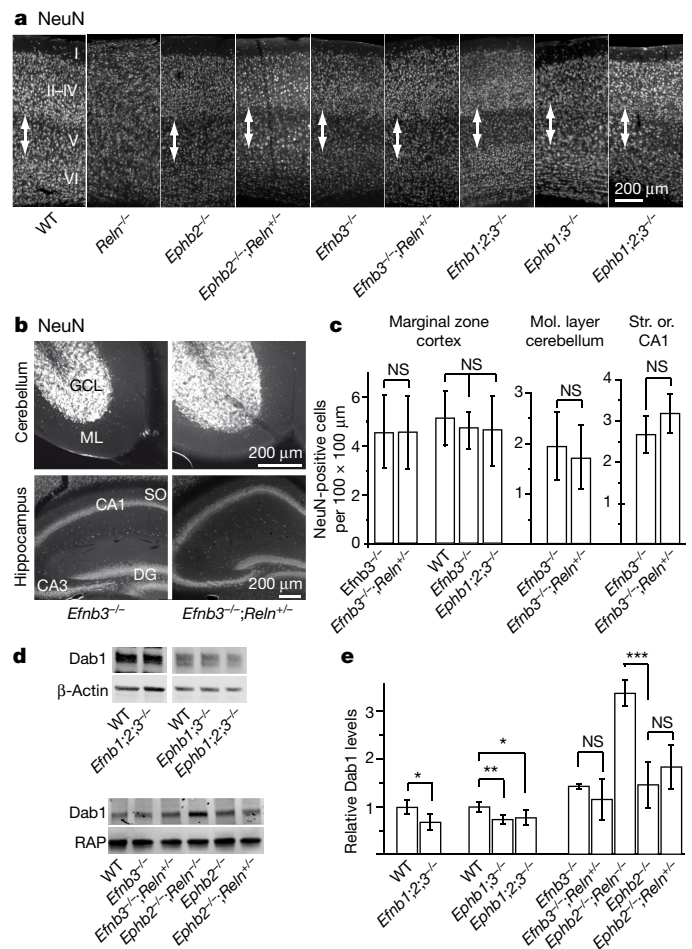


Figure 2 | *Efnb3*^{-/-}; *Reln*^{+/-} mice lack typical features of *reeler* mice such as over-migration of pyramidal neurons and elevated Dab1 levels. **a, b**, NeuN in the somatosensory cortex (**a**), cerebellum and hippocampus (**b**). Distinct layers I and V (double-headed arrows) are apparent, except for in *Reln*^{-/-} mice. **c**, Quantification of NeuN labelled neurons in different layers as indicated. Source data for the graphical representations are provided in the Supplementary Information. **d, e**, Dab1 levels (whole brain) did not differ significantly between *Efnb3*^{-/-} and *Efnb3*^{-/-}; *Reln*^{+/-} or *Ephb2*^{-/-} and *Ephb2*^{-/-}; *Reln*^{+/-} mice. In *Efnb1;2;3*^{-/-} (**P* < 0.05), *Ephb1;3*^{-/-} (***P* < 0.005), or *Ephb1;2;3*^{-/-} (**P* < 0.05) mice, Dab1 was slightly decreased, and increased threefold in *Ephb2*^{-/-}; *Reln*^{+/-} (***P* < 0.0005) mice; NS, not significant. Two-tailed Student's *t*-test. For western blot source data, see Supplementary Fig. 2. CA1, CA3, hippocampal areas; DG, dentate gyrus; GCL, dentate granular; ML, molecular layer; SO, strata oris; WT, wild type.

Although it is possible that Sentürk *et al.*² saw phenotypic variations in their *Efnb3*^{-/-}; *Reln*^{+/-} and *Efnb1;2;3*^{-/-} mice that depend on genetic background, this would merely reaffirm that *Efnb* ligands are in fact not essential for *Reln* pathway activation.

Our findings show that *Efnb* ligands and EphB receptors are not required for *Reln*-guided primary radial migration in the neocortex, whereas possible roles in pyramidal neuron axonal or dendritic growth, pruning, synaptogenesis and/or synaptic plasticity remain to be determined.

Methods

The source of mutant mouse lines, their breeding, as well as immunohistochemical and biochemical techniques, are described in the Supplementary Information.

Theresa Pohlkamp¹, Lei Xiao^{2,3}, Rukhsana Sultana^{2,3}, Asim Bepari^{2,3}, Hans H. Bock⁴, Mark Henkemeyer^{2,3} & Joachim Herz^{1,2,5,6}

¹Department of Molecular Genetics, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA.
email: joachim.herz@utsouthwestern.edu

²Department of Neuroscience, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA.

³Department of Developmental Biology, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA.

⁴Department of Gastroenterology, Hepatology and Infectious Diseases, Heinrich Heine University, Düsseldorf 40225, Germany.

⁵Department of Neurology and Neurotherapeutics, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA.

⁶Center for Translational Neurodegeneration Research, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA.

Received 7 November 2015; accepted 20 September 2016.

1. Rakic, P. & Caviness, V. S., Jr. Cortical development: view from neurological mutants two decades later. *Neuron* **14**, 1101–1104 (1995).
2. Sentürk, A., Pfennig, S., Weiss, A., Burk, K. & Acker-Palmer, A. Ephrin Bs are essential components of the Reelin pathway to regulate neuronal migration. *Nature* **472**, 356–360 (2011).
3. Xu, N. J. & Henkemeyer, M. Ephrin reverse signaling in axon guidance and synaptogenesis. *Semin. Cell Dev. Biol.* **23**, 58–64 (2012).
4. Bouché, E. *et al.* Reelin induces EphB activation. *Cell Res.* **23**, 473–490 (2013).
5. Yokoyama, N. *et al.* Forward signaling mediated by ephrin-B3 prevents contralateral corticospinal axons from recrossing the spinal cord midline. *Neuron* **29**, 85–97 (2001).
6. Villar-Cerviño, V. *et al.* Contact repulsion controls the dispersion and final distribution of Cajal–Retzius cells. *Neuron* **77**, 457–471 (2013).

7. Dahlstrand, J., Lardelli, M. & Lendahl, U. Nestin mRNA expression correlates with the central nervous system progenitor cell state in many, but not all, regions of developing central nervous system. *Brain Res. Dev. Brain Res.* **84**, 109–129 (1995).
8. Insolera, R., Bazzi, H., Shao, W., Anderson, K. V. & Shi, S. H. Cortical neurogenesis in the absence of centrioles. *Nat. Neurosci.* **17**, 1528–1535 (2014).
9. Thompson, C. L. *et al.* A high-resolution spatiotemporal atlas of gene expression of the developing mouse brain. *Neuron* **83**, 309–323 (2014).
10. Xu, N. J. & Henkemeyer, M. Ephrin-B3 reverse signaling through Grb4 and cytoskeletal regulators mediates axon pruning. *Nat. Neurosci.* **12**, 268–276 (2009).
11. Catchpole, T. & Henkemeyer, M. EphB2 tyrosine kinase-dependent forward signaling in migration of neuronal progenitors that populate and form a distinct region of the dentate niche. *J. Neurosci.* **31**, 11472–11483 (2011).
12. Kullander, K. *et al.* Ephrin-B3 is the midline barrier that prevents corticospinal tract axons from recrossing, allowing for unilateral motor control. *Genes Dev.* **15**, 877–888 (2001).
13. Xu, N. J., Sun, S., Gibson, J. R. & Henkemeyer, M. A dual shaping mechanism for postsynaptic ephrin-B3 as a receptor that sculpts dendrites and synapses. *Nat. Neurosci.* **14**, 1421–1429 (2011).
14. Grunwald, I. C. *et al.* Hippocampal plasticity requires postsynaptic ephrinBs. *Nat. Neurosci.* **7**, 33–40 (2004).
15. Rodenas-Ruano, A., Perez-Pinzon, M. A., Green, E. J., Henkemeyer, M. & Liebl, D. J. Distinct roles for ephrinB3 in the formation and function of hippocampal synapses. *Dev. Biol.* **292**, 34–45 (2006).

Supplementary Information is available in the online version of the paper.

Author Contributions T.P., L.X., R.S. and A.B. performed experiments. H.H.B., M.H. and J.H. designed the experiments. T.P. and J.H. wrote the manuscript.

Competing Financial Interests Declared none.

doi:10.1038/nature20129

Sentürk *et al.* reply

REPLYING TO T. Pohlkamp *et al.* *Nature* **539**, <http://dx.doi.org/10.1038/nature20129> (2016)

In 2011, our laboratory reported that cortical pyramidal neurons fail to migrate properly in *Efnb1;2;3*^{−/−} triple knockouts and in *Efnb3*^{−/−}; *Reln*^{+/−} compound mice¹. We provided genetic and biochemical evidence for an interaction between ephrin B (*Efnb*) ligands and Reelin (*Reln*) signalling during the development of laminated structures in the brain, such as the cortex, hippocampus and cerebellum. In the accompanying Comment², Pohlkamp *et al.* now revisit this question; however, they use a different *Efnb3* knockout strategy.

Although the authors² confirm our observation¹ that *Efnb* ligands and *Reln* interact genetically *in vivo*, as indicated by the increased *in utero* mortality of double knockout *Efnb3*^{−/−}; *Reln*^{−/−} embryos, they report their inability to reproduce a subset of the migration phenotypes originally reported by us. One likely explanation for this apparent discrepancy is the difference in the *Efnb3* knockout targeting strategy used in the two studies^{1,2}. The *Efnb3* allele used by us has a deletion in a large part of the first exon³ and was confirmed null by Southern blot analysis (see figure 1 of ref. 3). Pohlkamp *et al.*² used the mutant *Efnb3*^{neo/neo} in their analysis, which has been previously described by these authors to be hypomorphic (that is, a partial loss-of-function allele)⁴. Therefore, the possibility that residual *Efnb3* activity could rescue the specific subset of cortical migration phenotypes they fail to reproduce in their model cannot be excluded. Previous data stating that the *Efnb3*^{neo/neo} does not reproduce the long-term potentiation defects observed in our complete loss-of-function allele for *Efnb3* strongly indicates that this is the case^{5,6}. Moreover, the two alleles have also been reported to differ

in their effects on synapse formation⁷. Unfortunately, Pohlkamp *et al.*² cite neither this evidence nor their own previous work^{5,7}.

In addition to the possible effects caused by differing allele strengths, any differences in the penetrance of the phenotypes may also be partially attributed to differences in the genetic background of the mouse strains used. In this case, different sensitivity to *Reln* gene dosage in *Efnb3*^{−/−}; *Reln*^{+/−} compounds might be compensated by *Efnb*-independent mechanisms, as already proposed in our original manuscript¹.

Most notably, the complete absence or low quality of analysis performed by Pohlkamp *et al.*² (for example, they perform no quantification, show single examples of low-magnification images and use poorly stained samples with high background noise) should prevent any meaningful conclusions from being drawn and does not support the strong statements made by the authors. As a further example of this, Pohlkamp *et al.*² disregard some of the phenotypes that are apparent in their own representative images—they do not refer to the ectopic placement of neurons in the upper layers of the cortex in the *Efnb3*^{−/−}; *Reln*^{−/−} compound mice (as can be seen when comparing figure 1a and 1b of their Comment²) or to the lack of projections from CA1 pyramidal neurons to the stratum lacunosum-moleculare (see, among others, figure 1h of their Comment²).

It should be noted that, since publication of our paper¹, crosstalk between *Reln* and the EphB–*Efnb* system has been confirmed both biochemically⁸ and *in vivo*⁹, notably also by the authors of the

accompanying Comment². We certainly agree that advancing the field by uncovering the mechanisms by which these two important signalling pathways also interact in the context of processes such as neurite outgrowth, synaptogenesis and synaptic plasticity, remains an exciting subject for future investigations.

A. Sentürk¹, S. Pfennig¹, A. Weiss¹, K. Burk¹ & A. Acker-Palmer¹

¹Buchmann Institute for Molecular Life Sciences (BMLS) and Institute of Cell Biology and Neuroscience, Goethe University Frankfurt, Max-von-Laue-Str. 15, D-60438, Frankfurt am Main, Germany.
email: acker-palmer@bio.uni-frankfurt.de

1. Sentürk, A., Pfennig, S., Weiss, A., Burk, K. & Acker-Palmer, A. Ephrin Bs are essential components of the Reelin pathway to regulate neuronal migration. *Nature* **472**, 356–360 (2011).
2. Pohlkamp, T. *et al.* Ephrin Bs and canonical Reelin signalling. *Nature* <http://dx.doi.org/10.1038/nature20129> (2016).
3. Kullander, K. *et al.* Ephrin-B3 is the midline barrier that prevents corticospinal tract axons from recrossing, allowing for unilateral motor control. *Genes Dev.* **15**, 877–888 (2001).
4. Yokoyama, N. *et al.* Forward signaling mediated by ephrin-B3 prevents contralateral corticospinal axons from recrossing the spinal cord midline. *Neuron* **29**, 85–97 (2001).
5. Armstrong, J. N. *et al.* B-ephrin reverse signaling is required for NMDA-independent long-term potentiation of mossy fibers in the hippocampus. *J. Neurosci.* **26**, 3474–3481 (2006).
6. Grunwald, I. C. *et al.* Hippocampal plasticity requires postsynaptic ephrinBs. *Nat. Neurosci.* **7**, 33–40 (2004).
7. Aoto, J. *et al.* Postsynaptic ephrinB3 promotes shaft glutamatergic synapse formation. *J. Neurosci.* **27**, 7508–7519 (2007).
8. Bouché, E. *et al.* Reelin induces EphB activation. *Cell Res.* **23**, 473–490 (2013).
9. Catchpole, T. & Henkemeyer, M. EphB2 tyrosine kinase-dependent forward signaling in migration of neuronal progenitors that populate and form a distinct region of the dentate niche. *J. Neurosci.* **31**, 11472–11483 (2011).

doi:10.1038/nature20130

Contesting the massacre at Nataruk

ARISING FROM M. Mirazón Lahr *et al.* *Nature* **529**, 394–398 (2016); doi:10.1038/nature16477

Mirazón Lahr *et al.*¹ present the earliest evidence of inter-group warfare at the East African site of Nataruk. Their evidence of warfare is based on three inferences: that the skeletons were all contemporaneous, that the bodies were left unburied, and that most individuals exhibited perimortem trauma consistent with interpersonal violence. We believe the data suggest that the burials are not contemporaneous and that most of the observed cranial damage is inconsistent with blunt force trauma. Therefore, the inference of inter-group warfare is premature. There is a Reply to this Comment by Mirazón Lahr, M. *et al.* *Nature* **539**, <http://dx.doi.org/10.1038/nature19779> (2016).

Mirazón Lahr *et al.*¹ suggest that the skeletons at Nataruk document a single-event massacre of a forager band, after which the bodies remained unburied near the edge of a lagoon. The case for warfare is based on observations of alleged perimortem trauma to infer multiple simultaneous fatalities. Although the embedded obsidian lithic in KNM-WT 71251 is suggestive, interpretations of perimortem blunt force trauma on most crania are inconsistent with the forensic

anthropology literature^{2–6}. We discuss three of the more compelling examples here.

Mirazón Lahr *et al.*¹ describe a radiating fracture on the left frontal of KNM-WT 71253 as produced by an inferred strike to the left parieto-temporal region that was unobservable owing to erosion of the posterior vault. A similar fracture is noted on the right frontal. Both fractures are transversely oriented midway between the coronal suture and the supra-orbital margins and, importantly, are unconnected. Given the lack of an impact site, the right frontal fracture must be taphonomic, providing strong evidence for the presence of conditions conducive to post-depositional warping and cracking. This interpretation is supported by the directional change exhibited by this fracture—a feature shared, to a lesser extent, by the fracture on the left frontal described as resulting from blunt force. Such changes in direction have been associated experimentally with postmortem weathering processes⁷ (Fig. 1a, b) and fractures with this morphology are not interpreted as evidence for trauma in the bioarchaeological literature^{8–11}.

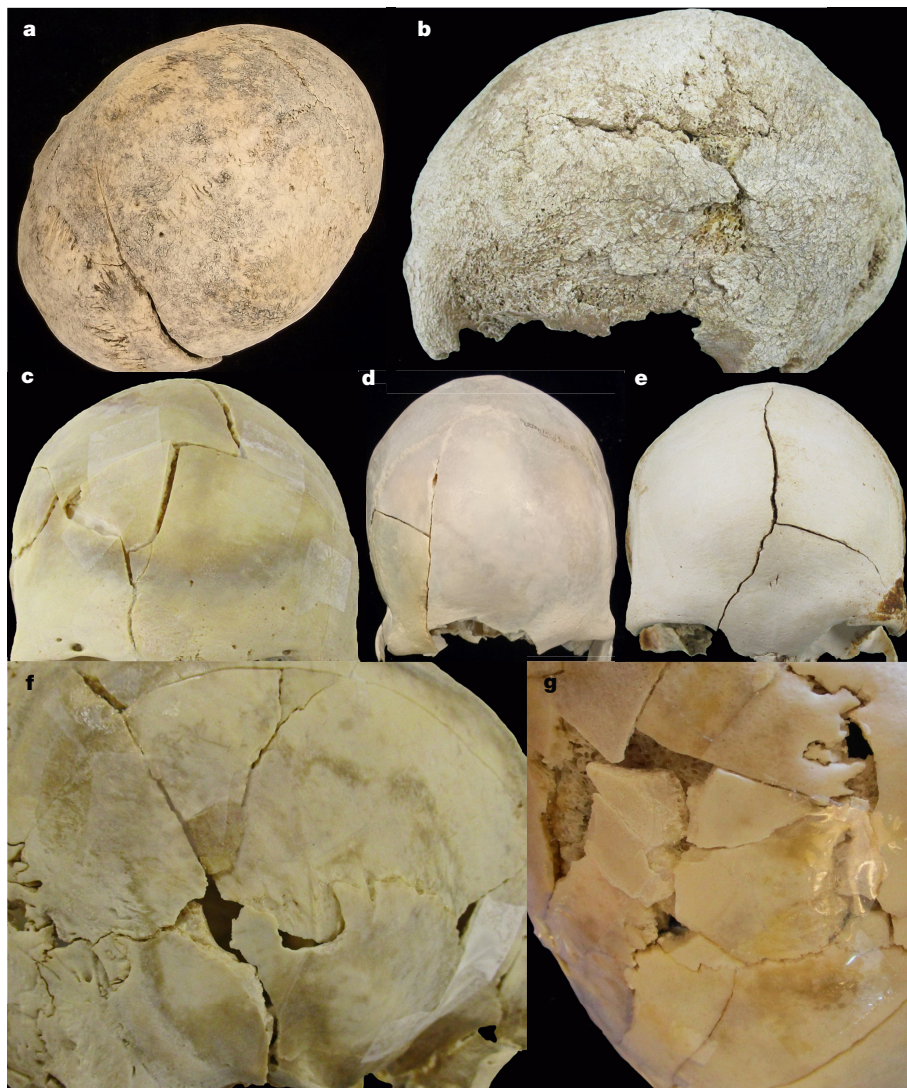


Figure 1 | Examples of cranial vault taphonomy and trauma from modern forensics cases. a, b, Radiating fractures resulting from exposure and weathering, note the similar appearance of 1A to KNM-WT 71257, and the similar appearance of 1B to KNM-WT 71251; c–e, and inferior portion of g, vertically oriented fracture lines resulting from blunt force trauma to the frontal bone; f, radiating and concentric fractures resulting from blunt force trauma to the right temporo-parietal with convergence of radiating fractures at impact point; g, wedge-shaped fragments resulting from repeated blows to the right parietal, note chipping of fracture edges, separation of internal and external tables, and absence of fractures oriented in a grid-like pattern.

Given these observations and the erosion of the posterior cranial vault, we believe that the left frontal fracture is also taphonomic in origin.

KNM-WT 71264 shows a large horizontal fracture on the frontal bone interpreted as evidence of blunt force trauma. However, the fracture pattern is inconsistent with the expectations of a direct impact to this region and the mechanics of fracture propagation. The morphology of the fracture is irregular, exhibiting multiple abrupt changes in direction with a transverse orientation that crosses the midline. Numerous studies indicate that blunt force trauma to the frontal squama results in vertically-oriented fractures due to the midfrontal buttressing of the cranial vault^{2,3,12} (Fig. 1c–e, g). Fractures approaching such buttresses obliquely, as in the images provided by Mirazón Lahr *et al.*¹, typically divert towards areas that are less structurally reinforced³. This is inconsistent with the pattern observed in KNM-WT 71264. When fractures of the frontal bone are oriented transversely, they tend to be located more inferiorly and traverse the area of weakness associated with the frontal sinus, which is conspicuously intact in the images provided by Mirazón Lahr *et al.*¹. In a large sample of experimental ($n = 195$) and clinical ($n = 290$) cases of frontobasal fractures, Manson and colleagues¹² recorded only three different patterns of fracture propagation on the frontal. None of these match the appearance of KNM-WT 71264.

For KNM-WT 71257, Mirazón Lahr *et al.*¹ describe an impact to the left temporal causing concentric and radiating vault fractures, basicranial displacement, and a radiating fracture affecting the right side of the facial skeleton. Focal blunt force trauma to the cranial vault typically results in depression at the point of impact, producing localized in-bending of the vault with consequent out-bending in the surrounding area. As these processes progress, radiating fractures propagate outward from the point of impact and concentric fractures may form perpendicular to the radiating fractures, encircling the impact site. This configuration of fractures typically results in wedge-shaped pieces of bone^{2–4,6} (Fig. 1f, g). This is not seen in the images provided by Mirazón Lahr *et al.*¹. Rather, breakage of the left temporal region has a tessellated appearance resulting from a series of parallel vertically oriented fractures intersected at nearly right angles by several horizontally oriented fractures. These fracture lines have no single origin that would represent the site of impact. Note also that large portions of the temporal squama and the greater wing of the sphenoid appear to be intact. Given the thinness of these bones, it is unlikely that a direct blow impacted this region. Although Mirazón Lahr *et al.*¹ indicate that damage to this region was exacerbated by soil compression, the observable configuration of fracture lines is inconsistent with what is expected to result from blunt force trauma to the temporal region. The same argument applies to the inferred blunt force trauma to the left temporal bone of KNM-WT 71264.

Perimortem trauma is difficult to diagnose in ancient remains with complex taphonomic histories. Mirazón Lahr *et al.*¹ base their interpretations on six criteria, but their assessment of these in the Nataruk skulls is open to question (see above), and several criteria may have limited utility in discriminating perimortem trauma from postmortem damage^{13,14}. We suggest soil compression is the primary cause of damage in KNM-WT 71253, 71264 and 71257. That the burials were found in lakebed sediments indicates they were submerged in antiquity and may have decomposed under conditions that increased bone malleability through mineral leaching³. This interpretation is supported by an association between *in situ* skull orientation and the location of primary cranial vault damage and contour disruption (Table 1). Individuals buried face down show primary damage in the antero-posterior direction, while individuals buried with their skulls on their side show primary damage in the medio-lateral direction. This is unlikely to be a coincidence. Although difficult to diagnose from the

Table 1 | Body positioning and dating of the 12 *in situ* Nataruk burials

Burial number	Body position	Skull orientation	Uranium series date ^a	¹⁴ C shell age (cal BP)	Primary location of vault damage
71251	prone	face down	12,200 ± 700	10,275–10,190	Anterior
71253	prone ^b	face down	12,200 ± 730 9,540 ± 860		Anterior
71264	prone	face down	48,300 ± 2900	10,480–10,470 10,420–10,230	Anterior, lateral
71263	side/ prone ^c	partly down			Anterior ^d
71265		partly down ^e			Anterior
71254	side	on side	6,640 ± 460	9,290–9,030	Medio-lateral
71257	side	on side	5,960 ± 550 6,670 ± 510		Medio-lateral
71258	side			9,460–9,260	
71259	seated	face down ^f	5,590 ± 450 6,590 ± 740		
71255	seated	face up ^g			
71256	supine	face up			
71260	side	on side			

^aOnly uranium series dates that do not violate the diffusion-adsorption-decay model assumptions are included here. Data for both columns taken directly from the 'Age' column in extended data table 1 in Mirazón Lahr *et al.*¹.

^bTorso and skull are prone with knees flexed to the chest.

^cIt is difficult to reconstruct the burial position of this individual based on the images provided in Mirazón Lahr *et al.*¹. Figure 2j in ref. 1 clearly shows a prone posture; however, this is difficult to reconcile with the composite image presented in Fig. 1 and extended data figure 7a in ref. 1.

^dFractures do not fit within scoring scheme as there is no real vault deformation, face is absent with primary inferred trauma to the frontal.

^eMirazón Lahr *et al.*¹ describe this 'as the face was partly turned down into the ground'.

^fMirazón Lahr *et al.*¹ describe the skull orientation as face down, but also state the cranium was exposed and not recovered.

^gMirazón Lahr *et al.*¹ infer this based on the orientation of the cervical vertebrae.

images provided, inferred postcranial damage is also consistent with taphonomic changes caused by wet-dry cycling, soil compression, and root, animal and insect activity; none are given serious consideration in a differential diagnosis by Mirazón Lahr *et al.*¹.

Furthermore, although interpreted conservatively as minimum age estimates, uranium series dates are associated with burial posture (prone versus non-prone). Importantly, the same pattern holds for radiocarbon dates based on shell ($P = 0.012$), which are not subject to the same interpretive limitations. The oldest dates are recorded for prone burials and the youngest dates associate with individuals interred on their side (Table 1). This pattern suggests a multi-component chronology consistent with mortuary sequences in Sudan where multiple prone inhumations have been identified in attritional cemetery contexts¹⁵.

Finally, it is unlikely that the Nataruk skeletons were unburied and deposited simultaneously as suggested by Mirazón Lahr *et al.*¹. Their interpretation is based on three observations: (1) the lack of burial pits, (2) variation in body position, and (3) the association between fish bones and human remains. The absence of burial pits is commonplace in mortuary sites throughout the world, especially of this age. Unlike well-documented massacre sites, the bodies at Nataruk are mostly articulated, spatially organized, non-commingled, and preserve limited variation in body positioning for which clear categories can be defined (see Table 1). Their posited non-burial scenario is not the only way for shell and fish bone to adhere to human bone within the post-depositional environment. The interpretation of Mirazón Lahr *et al.*¹ requires the decomposition of dozens of bodies at the edge of a lagoon,

without evidence of disarticulation or animal scavenging, long enough for them to be covered by lakebed sediments in an ephemeral and low energy depositional environment. This is unlikely, and in the case of the seated burial (71255) not possible. Interpersonal violence was surely present in early Holocene African hunter-gatherers, however, the case for a massacre at Nataruk is not supported by the data Mirazón Lahr *et al.*¹ report.

Methods

Data for Table 1 were extracted from Mirazón Lahr *et al.*¹. We only include uranium series dates that did not violate the diffusion-adsorption-decay model assumptions. Damage location is based on the anatomical plane of the cranial vault that shows the most significant displacement in its contour. Assessments of blunt force cranial trauma are based on images and descriptions of fractures. Figure 1 was created using modern forensic cases from Maricopa County, Arizona as exemplars of typical patterning resulting from cranial blunt force traumata. *t*-test is based on the uncalibrated midpoint of radiocarbon dates using a two-sided test and the Satterthwaite approximation.

**Christopher M. Stojanowski^{1,2}, Andrew C. Seidel^{1,2},
Laura C. Fulginiti^{1,3}, Kent M. Johnson^{1,2} & Jane E. Buikstra^{1,2}**

¹School of Human Evolution and Social Change, Arizona State University, Tempe, Arizona 85287, USA.

email: cstojano@asu.edu

²Center for Bioarchaeological Research, Arizona State University, Tempe, Arizona 85287, USA.

³Maricopa County Office of the Medical Examiner, Phoenix, Arizona 85007, USA.

Received 2 March; accepted 9 August 2016.

1. Mirazón Lahr, M. *et al.* Inter-group violence among early Holocene hunter-gatherers of West Turkana, Kenya. *Nature* **529**, 394–398 (2016).

2. Berryman, H. E. & Symes, S. A. in *Forensic Osteology* (ed. Reichs, K. J.) 333–352 (Charles C. Thomas, 1998).
3. Galloway, A. & Wedel, V. L. in *Broken Bones* (ed. Wedel, V. L. & Galloway, A.) 133–160 (Charles C. Thomas, 2014).
4. Kroman, A., Kress, T. & Porta, D. Fracture propagation in the human cranium: a re-testing of popular theories. *Clin. Anat.* **24**, 309–318 (2011).
5. Sauer, N. J. in *Forensic Osteology* (ed. Reichs, K. J.) 321–332 (Charles C. Thomas, 1998).
6. Symes, S. A., L'Abbé, E. N., Chapman, E. N., Wolff, I. & Dirkmaat, D. C. in *A Companion to Forensic Anthropology* (ed. Dirkmaat, D. C.) 340–389 (Wiley-Blackwell, 2012).
7. Calce, S. & Rogers, T. L. Taphonomic changes to blunt force trauma: a preliminary study. *J. Forensic Sci.* **52**, 519–527 (2007).
8. Crist, T. A. J., Washburn, A., Park, H., Hood, I. & Hickey, M. A. in *Forensic Taphonomy* (ed. Haglund, W. D. & Sorg, M. H.) 319–336 (CRC Press, 1997).
9. Erdal, Ö. D. A possible massacre at early Bronze Age Titriş Höyük, Anatolia. *Int. J. Osteoarchaeol.* **22**, 1–21 (2012).
10. Gheggi, M. S. Conflict in pre-Hispanic northwest Argentina: implications arising from human bone trauma patterns. *Int. J. Osteoarchaeol.* **26**, 17–27 (2016).
11. Scott, R. M. & Buckley, H. R. Biocultural interpretations of trauma in two prehistoric Pacific Island populations from Papua New Guinea and the Solomon Islands. *Am. J. Phys. Anthropol.* **142**, 509–518 (2010).
12. Manson, P. N. *et al.* Frontobasal fractures: anatomical classification and clinical significance. *Plast. Reconstr. Surg.* **124**, 2096–2106 (2009).
13. Zephro, L. *Determining the Timing and Mechanism of Bone Fracture in Bovine Bone*. Ph.D. Diss., Dept. of Anthropology, University of California, Santa Cruz. (2012).
14. Galloway, A., Zephro, L. & Wedel, V. L. in *Broken Bones* (ed. Wedel, V. L. & Galloway, A.) 47–58 (Charles C. Thomas, 2014).
15. Salvatori, S., Usai, D. & Zerboni, A. Mesolithic site formation and palaeoenvironment along the White Nile (central Sudan). *Afr. Archaeol. Rev.* **28**, 177–211 (2011).

Author Contributions Principal writing of the paper by C.M.S. and A.C.S., with contributions from L.C.F., K.M.J., and J.E.B. C.M.S. and K.M.J. were responsible for site interpretations. A.C.S. and L.C.F. were responsible for fracture pattern interpretations.

Competing Financial Interests Declared none.

doi:10.1038/nature19778

Mirazón Lahr *et al.* reply

REPLYING TO C. M. Stojanowski *et al.* *Nature* **539**, <http://dx.doi.org/10.1038/nature19778> (2016)

In the accompanying Comment¹, Stojanowski *et al.* challenge the evidence for inter-group conflict at Nataruk². They make two arguments—first, that the lesions in three crania are due to soil compression; second, that there is a correlation between body position and age, reflecting different burial traditions. We believe that their interpretation is incorrect on both counts.

Stojanowski *et al.*¹ argue that the fractures in three Nataruk skulls do not meet forensic criteria, the criteria are unreliable, and the location of lesions is linked to skull/face orientation. We focus on cranium 71264, argued by Stojanowski *et al.*¹ as an example of soil compression effects. This skull has two primary lesions (challenging any site-wide association between skull trauma location and face orientation), one on the left temporoparietal area, another on the frontal bone; these show clear in-bending of bone at impact-point and out-bending of affected regions, radiating fractures, directional change, depressed adhering bone fragments, lesions across sinuses, evulsion of molars, and differential lateral displacement of maxilla and mandible. We know of no selectively targeted, multi-directional soil compression agent that could account for the observed pattern in this, or other Nataruk crania. The criteria used for identification of perimortem trauma are not only consistent with the literature^{3,4}, but clearly identify multiple traumatic lesions at Nataruk².

Stojanowski *et al.*¹ also argue that prone individuals at Nataruk are older than the rest, and that the site must be a cemetery, with sequential burial traditions, to have preserved the remains. To support this, they (1) claim that there are four prone individuals at Nataruk, not two, by redefining the word ‘prone’ thus including two skeletons found flexed on their side²; (2) dismiss the absence of burial pits or the shells/fish adhering onto the bones as irrelevant; (3) do not account for the fact that bodies and body parts were strewn over 20,000 m²; (4) incorrectly state that the majority of human remains at Nataruk were articulated and uncommingled; (5) argue for ‘limited variation in body positioning’ (figure 1 from ref. 2 shows clearly the diversity of position/direction of bodies, limbs, heads and faces at Nataruk); and (6) draw parallels with cemeteries in Sudan⁵. These are rich multi-stratified, younger sites (~8,000–2,000 years BP), in one of which earlier burials (with residual burial pits) were prone, later Neolithic ones, flexed³. However, prone skeletons in Sudan, although rare, are found across time—from 30,000–20,000 years ago at Wadi Kubaniya⁶ to as recent as 1st millennium BC at Jebel Moya^{7,8}, and preserved unburied bodies found before in East Africa⁹. In fact, there are no chronological, stratigraphic or internment parallels between the Omdurman sites and Nataruk, as there are no universal East African pre-Neolithic burial traditions.

BRIEF COMMUNICATIONS ARISING

Nataruk is not a cemetery, the majority of remains were disarticulated and dispersed; the human remains' multiple cranial and post-cranial traumas are more consistent with perimortem traumatic lesions than soil compression, and critically include embedded projectiles. There is no association between body position and age, or face orientation and location of fractures. A case of inter-group conflict remains the best explanation of the events at Nataruk.

M. Mirazón Lahr^{1,2}, F. Rivera¹, R. K. Power¹, A. Mounier¹, B. Copsey¹, F. Crivellaro¹, J. E. Edung³, J. M. Maillo Fernandez⁴, C. Kiarie², J. Lawrence¹, A. Leakey², E. Mbua⁵, H. Miller¹, A. Muigai⁶, D. M. Mukhongo¹, A. Van Baelen¹, R. Wood⁷, J.-L. Schwenninger⁸, R. Grün^{7,9}, H. Achyuthan¹⁰, A. Wilshaw¹ & R. A. Foley^{1,2}

¹Leverhulme Centre for Human Evolutionary Studies, Department of Archaeology & Anthropology, University of Cambridge, Fitzwilliam Street, Cambridge CB2 1QH, UK.

email: mbml1@cam.ac.uk

²Turkana Basin Institute, PO box 24467-00502, Nairobi, Kenya.

³National Museums of Kenya, Jomo Kenyatta House, PO box 152-30500, Lodwar, Kenya.

⁴Departamento de Prehistoria y Arqueología, UNED, c/ Paseo Senda del Rey, 7, 28040 Madrid, Spain.

⁵National Museums of Kenya, PO box 40658-00100, Nairobi, Kenya.

⁶Jomo Kenyatta University of Agriculture and Technology, PO box 62000-00200, Nairobi, Kenya.

⁷Research School of Earth Sciences, The Australian National University, Building 142, Mills Road, Acton, Australian Capital Territory 2601, Australia.

⁸Research Laboratory for Archaeology and the History of Art, University of Oxford, Dyson Perrins Building, South Parks Road, Oxford OX1 3QY, UK.

⁹Australian Research Centre for Human Evolution, Environmental Futures Research Institute, Griffith University, 170 Kessels Road, Nathan, Queensland 4111, Australia.

¹⁰Department of Geology, Anna University, Chennai, Tamil Nadu 600025, India.

1. Stojanowski, C. M. *et al.* Contesting the massacre at Nataruk. *Nature* **539**, <http://dx.doi.org/10.1038/nature19778> (2016).
2. Mirazón Lahr, M. *et al.* Inter-group violence among early Holocene hunter-gatherers of West Turkana, Kenya. *Nature* **529**, 394–398 (2016).
3. Buikstra, J. E. & Ubelaker, D. H. Standards for data collection from human skeletal remains. *Am J Hum Biol.* **44**, 120 (1994).
4. Knüsel, C. & Smith, M. J. (Eds). *The Handbook of the Bioarchaeology of Human Conflict*. (Routledge, 2013).
5. Salvatori, S., Usai, D. & Zerboni, A. Mesolithic site formation and palaeoenvironment along the White Nile (Central Sudan). *Afr. Archaeol. Rev.* **28**, 177–211 (2011).
6. Wendorf, F. & Schild, R. *The Wadi Kubbaniya skeleton: A Late Paleolithic burial from southern Egypt* 71–74 (Southern Methodist University Press, 1986).
7. Addison, F. *Jebel Moya* (Oxford University Press, 1949).
8. Brass, M. & Schwenninger, J. L. Jebel Moya (Sudan): new dates from a mortuary complex at the southern Meroitic frontier. *Azania* **48**, 455–472 (2013).
9. Leakey, L. S. B. The Naivasha fossil skull and skeleton. *J. East African Natural History Society* **16**, 169–177 (1942).

doi:10.1038/nature19779

DNA REPAIR

Clamping down on copy errors

Repair enzymes must communicate across hundreds of nucleotides to undo errors made during DNA replication. Imaging reveals that the enzymes do this by forming a series of ring-like clamps that diffuse along the DNA. [SEE LETTER P.583](#)

NEIL M. KAD & BENNETT VAN HOUTEN

As we write up results for publication, any spelling errors are quickly identified by our computer's built-in spell-checker. However, this is not an invention of the modern world; nature has been using DNA spellcheckers for millions of years to avoid genetic errors that arise during DNA replication, which it corrects through a process called mismatch repair. In the bacterium *Escherichia coli*, repair requires communication between enzymes across long stretches of DNA, and how this occurs has been hotly debated for decades^{1–3}. On page 583, Liu *et al.*⁴ help to solve this mystery by using state-of-the-art techniques to analyse mismatch repair at the single-molecule level.

DNA exists as a double-stranded duplex, connected across the strands by complementary base-pairing — guanine (G) with cytosine (C) and adenine (A) with thymine (T). During DNA replication, the duplex unwinds, and each strand is copied using new nucleotides to create a freshly synthesized daughter strand. Errors occur when incorrect bases are incorporated into the daughter strand, creating a mismatch that requires subsequent repair to prevent mutations from arising. In *E. coli*, DNA with the base sequence GATC is normally tagged with a methyl group, but the newly synthesized DNA is temporarily unmethylated. These 'hemi-methylated' regions provide a means for mismatch-repair enzymes to distinguish between parent and daughter DNA.

Hemi-methylated sequences are also the sites at which the endonuclease enzyme MutH generates a single-strand break in the erroneous DNA strand during mismatch repair. Because of the rarity of GATC sequences, MutH-generated breaks can occur hundreds of bases from the site of the mismatch. Following breakage, the damaged strand is excised by an exonuclease enzyme that works from the break to the error, and the DNA is resynthesized to incorporate the correct base. A key aspect of this process is the coordination of mismatch detection with loading of the DNA-excision machinery, to ensure that excision occurs towards, rather than away from, the mismatch.

For the past decade, it has been thought^{5,6} that the protein MutS, which recognizes DNA mismatches, acts as a sliding clamp — a ring around DNA that can efficiently move up or down the two strands, enabling communication between the mismatch and the hemi-methylated sites. A third protein, MutL, activates the cutting activity of MutH. But how MutS links mismatch detection to the activation of MutH has been unclear.

To study initiation of mismatch repair, Liu *et al.* added fluorescent tags to MutH, MutS and MutL and analysed the movements of the individual proteins along single DNA molecules. They found that MutL is loaded onto DNA through interactions with a MutS molecule that has already located a mismatch. This observation builds on a structural study⁷ showing that the DNA duplex is positioned in the open channel of the MutS ring following mismatch verification, and that MutL

binding sites on MutS are revealed during this process, permitting formation of the MutS–MutL complex.

Next, the authors confirmed another previous observation⁶ — that the MutS–MutL complex forms a single sliding clamp. Extending these observations, they showed that MutL can detach from MutS to act as a clamp on its own. In this role, MutL undertakes short excursions away from MutS. Surprisingly, some of these excursions pass behind MutS, an observation that Liu *et al.* took to imply that MutL might pass through the MutS channel. Therefore, MutL can be thought of as acting like a yo-yo, separately scouting the DNA. But scouting for what?

The answer comes from the group's demonstration that MutL can recruit MutH to the DNA, thereby setting the stage for the excision phase of mismatch repair. The MutL–MutH complex behaves similarly to MutL alone, detaching from MutS and diffusing along DNA more quickly than when the three proteins are bound together — presumably in search of hemi-methylated sites. The authors propose that the MutS–MutL–MutH complex is the dominant species, because the sliding-clamp nature of MutS means that this complex is tethered to the DNA. However, efficient searching is also possible with MutL–MutH, which Liu and colleagues posit hops along the DNA.

Together, these observations might have solved the long-standing mystery of how the mismatch and excision sites interact, superseding previous models in which the

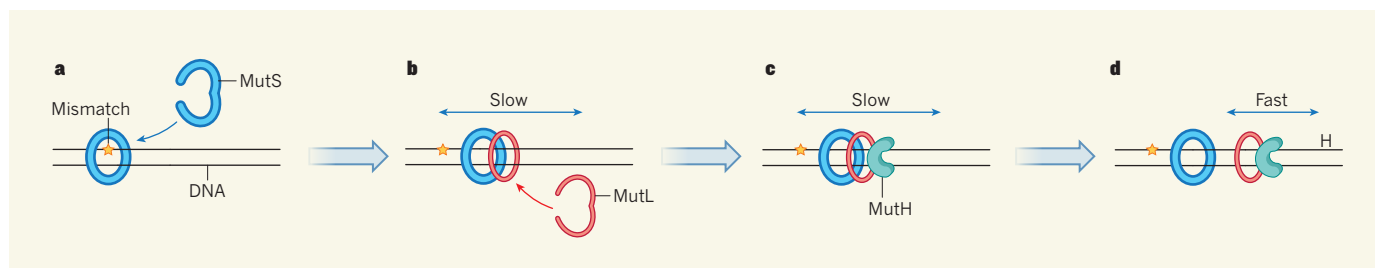


Figure 1 | Scouting for errors. **a**, Following DNA replication, the protein MutS clamps around DNA at mismatches — sites of incorrect pairing between the two complementary DNA strands. **b**, Liu *et al.*⁴ report that, once a mismatch is detected, MutS recruits the protein MutL, and the two proteins slowly slide together up and down DNA. **c**, MutL in turn recruits the enzyme

MutH, and the three move as one complex. **d**, MutL and MutH can detach from MutS to move rapidly along DNA in search of hemi-methylated sites (H) — sequences at which the parent strand is tagged with a methyl group but the mismatched daughter strand is not. The daughter strand is subsequently cleaved by MutH, initiating DNA excision and repair.

DNA is looped through the MutS clamp or in which the clamp moves in a targeted manner along the duplex^{1,8}. The model that emerges from these data suggests that MutS acts as a guiding clamp, sending out scouting clamps to search for hemi-methylated GATC sequences in a highly energy-efficient manner (Fig. 1).

In 2015, Paul Modrich received the Nobel Prize in Chemistry for his seminal contributions to our understanding of mismatch repair¹. However, many aspects of the process remain unresolved. One is that mismatch repair in eukaryotic cells (those that have a nucleus) involves more-complex protein–protein and protein–DNA interactions that have yet to be fully described. For instance, the eukaryotic equivalent of MutL has endonuclease activity, and eukaryotes use different mechanisms for excision depending on whether the break point is 3' or 5' of the mismatch^{1,3}. Given this complexity, a better understanding of eukaryotic mismatch repair will surely present new mechanistic surprises. The use of single-molecule approaches to enable analysis of the eukaryotic process, already under way in several laboratories, will be central to elucidating these systems.

The principles outlined by Liu *et al.* might apply to enzymes involved in other types of DNA repair. A recent report⁹ revealed that the yeast protein Rad4, which is involved in nucleotide-excision repair, does not directly bind damaged sites, but instead diffuses for up to 1 kilobase around the damage, providing a dynamic platform for the recruitment of other repair proteins. Furthermore, the protein Mfd, which dislodges RNA polymerase enzymes that become stalled while transcribing DNA, then scouts ahead for DNA damage on the transcribed strand¹⁰.

The next chapter of the mismatch-repair story will surely see experiments that follow the entire process of repair, enabling observation of every protein in real time in a single assay. With single-molecule fluorescence techniques developing rapidly, such experiments are in sight, and should provide a molecular understanding of biology *in vitro* that can be adapted to systems *in vivo*. ■

Neil M. Kad is in the School of Biosciences, University of Kent, Canterbury CT2 7NJ, UK. **Bennett Van Houten** is in the Department of Pharmacology and Chemical Biology, University of Pittsburgh, University of Pittsburgh Cancer Institute, Hillman Cancer Center, Pittsburgh, Pennsylvania 15213, USA.

6. Gorman, J. *et al.* *Proc. Natl Acad. Sci. USA* **109**, E3074–E3083 (2012).
7. Groothuizen, F. S. *et al.* *eLife* **4**, e06744 (2015).
8. Iyer, R. R., Pluciennik, A., Burdett, V. & Modrich, P. L. *Chem. Rev.* **106**, 302–323 (2006).

9. Kong, M. *et al.* *Mol. Cell* **64**, 376–387 (2016).
10. Haines, N. M., Kim, Y.-I. T., Smith, A. J. & Savery, N. J. *Proc. Natl Acad. Sci. USA* **111**, 4037–4042 (2014).

This article was published online on 16 November 2016.

MATERIALS SCIENCE

Polymers make charge flow easy

Organic semiconductor devices require good electrical contacts with conducting materials, but such contacts are often inefficient. An approach that tackles this problem will enable a wide range of applications. SEE LETTER P.536

ANTONIO FACCHETTI

Unconventional optoelectronic devices such as thin-film transistors, solar cells and light-emitting diodes, in which organic semiconductors are key components, will potentially give rise to new display, circuit, lighting, imaging and energy-harvesting products^{1–3}. All of these devices comprise a multilayer architecture in which a thin film of a conductor — typically, a metal or a transparent conducting oxide — is in contact with a semiconductor. To optimize the transfer of

electric charge between these two materials, an interfacial layer (IFL) is usually introduced, but the resulting architectures are often inefficient. On page 536, Tang *et al.*⁴ show that IFLs created from a particular type of polymer can improve both the performance and the stability of optoelectronic devices.

Several approaches have been developed to engineer IFLs, and these mainly depend on whether most of the charges that cross the conductor–semiconductor interface are electrons (negative charges) or holes (positive charges)⁵. For electron exchange, the energy needed to

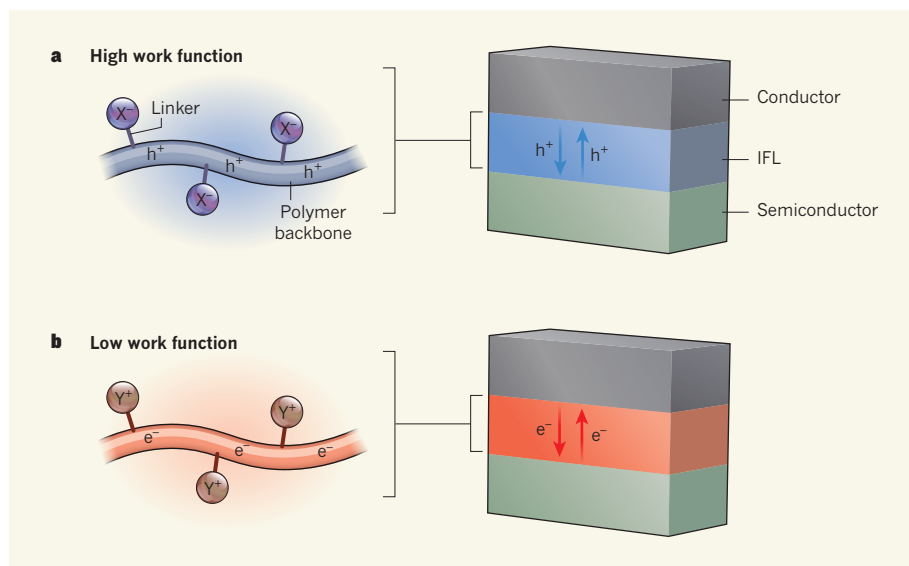


Figure 1 | From chemistry to optoelectronics. Tang *et al.*⁴ have developed a strategy to fabricate optoelectronic devices in which electrical charges can flow freely between the conductor and semiconductor components. **a**, To achieve this, the authors use organic semiconductors called self-compensated π -conjugated polymers as an interfacial layer (IFL) between the two materials. These polymers consist of 'backbones' to which positively charged holes (h^+) can be added, referred to as hole doping. The charges of these holes are balanced by negatively charged ions (X^-) that are chemically bonded to the polymer backbone by electrically insulating 'linkers'. Tang and colleagues show that hole doping such a polymer increases the energy that is needed to remove an electron from the conductor (the work function), allowing holes to be freely transferred through the IFL. **b**, Conversely, the authors show that by adding electrons (e^-) to the polymer backbone — electron doping — the charges of the electrons are balanced by positively charged ions (Y^+), and the system has a low work function, enabling electrons to be freely transferred through the IFL.

1. Modrich, P. *Angew. Chem. Int. Edn* **55**, 8490–8501 (2016).
2. Fishel, R. J. *Biol. Chem.* **290**, 26395–26403 (2015).
3. Kunkel, T. A. & Erie, D. A. *Annu. Rev. Genet.* **49**, 291–313 (2015).
4. Liu, J. *et al.* *Nature* **539**, 583–587 (2016).
5. Jeong, C. *et al.* *Nature Struct. Mol. Biol.* **18**, 379–385 (2011).

remove an electron from the conductor, called the work function, should be low and similar to that of the empty conduction energy band of the semiconductor. Conversely, for hole exchange, the work function should be high and match that of the semiconductor's filled valence band.

The most prominent families of materials used for IFLs include polyelectrolytes, metal oxides, self-assembled mono- or multilayers of organic molecules and mixtures of these^{6,7}. Each of these families has its strengths and limitations. However, using only one type of IFL family has made it difficult to secure useful properties, such as the ability to tune the work function over a broad range of energies, large conductivities, equally efficient hole and electron transfer and IFL stability.

Tang and colleagues have constructed IFLs using organic semiconductors called π -conjugated polymers, which allow efficient inter- and intramolecular charge transport to occur between the conductor and semiconductor materials⁸. Most π -conjugated polymers cannot transfer a large number of charged particles. However, the authors use a chemical doping process — the addition of either holes or electrons — to dramatically increase the density of charge carriers in the 'backbone' of these polymers and, in turn, their electrical conductivity (Fig. 1).

In conventional doped polymers, the charge of the delocalized carriers in the backbone is balanced by molecular counterions (ions that have the opposite charge to that of the carriers). However, Tang and collaborators' polymers are self-compensated — the balance of charge is obtained intramolecularly owing to counterions that are chemically bonded to the polymer backbone by electrically insulating linkers. Such polymers, particularly hole-doped ones, have long been studied⁹, and recent reports have extended the repertoire to electron-doped varieties¹⁰. The authors make the case for the unique properties of these polymers and their broad applicability in a vast array of optoelectronic devices.

In Tang and colleagues' polymers, doping and self-compensation are achieved in two separate steps, potentially improving control of the level of doping and the purity of the resulting materials. Furthermore, this control can be achieved both in solution and in thin films, broadening the range of doping materials and polymer structures that could be used for IFL fabrication. The array of self-compensated polymers includes hole-doped triarylamine-fluorene copolymers that can provide ultra-high work functions (up to 5.8 electronvolts) and electron-doped naphthalene-bis(carboxyimide)-fluorene copolymers that can have ultra-low work functions (down to 3.0 eV). Such a wide range of values allows the transfer of charge between the conductor and semiconductor materials to be optimized.

In every multilayer optoelectronic device,

chemical stability of the material components is of paramount importance for operational stability. This aspect is of even greater concern when using doped polymers, because ions can diffuse across the interfaces between layers, particularly in the presence of an electric field. However, Tang *et al.* show that, because the counterions in their self-compensated polymers are covalently bonded through the linkers, the diffusion of charge carriers and ions between layers is prevented. This molecular design is essential for achieving a stable concentration of the carriers (a constant doping density) across the entire thickness of the IFL film and across the interface between the IFL and the semiconductor.

Tang and colleagues fabricate several devices in which standard IFL materials are replaced by their polymers, including thin-film transistors, solar cells, light-emitting diodes and photodiodes. The authors' realization of a blue-light-emitting diode is particularly impressive. Here, they use two self-compensated polymers (one electron-doped and one hole-doped) and a benchmark organic blue-light emitter (a polyfluorene) to achieve electric currents and energy-conversion efficiencies that surpass those of conventional IFLs using PEDOT:PSS and calcium — materials that have high and low work functions, respectively.

The results demonstrate the great potential of self-compensated polymers, and raise several technological challenges and scientific questions. From a chemistry perspective, it will be instructive to study the synthetic reproducibility and scale-up feasibility of these polymers — to check whether the doping level and ionic content can still be controlled, and how charge transport and the work

function are affected. Furthermore, although the authors demonstrate the stability of their materials (and the work functions of these materials), they do not evaluate the stability of the devices incorporating their IFLs according to commercial standards. Such protocols are typically extremely demanding for use in commercial products.

Finally, an essential but unanswered question is whether self-compensation can result in electron-conducting polymers that have electrical conductivities and stabilities at ambient conditions that match those of current benchmark doped polymers such as PEDOT:PSS or polyaniline. However, despite the above-mentioned challenges, Tang and colleagues' work is a step on the path towards unconventional optoelectronics. ■

Antonio Facchetti is in the Department of Chemistry, Northwestern University, Evanston, Illinois 60208, USA, and at Polyera Corporation, Skokie, Illinois.
e-mail: a-facchetti@northwestern.edu

1. Yu, X., Marks, T. J. & Facchetti, A. *Nature Mater.* **15**, 383–396 (2016).
2. Vosgueritchian, M., Tok, J. B.-H. & Bao, Z. *Nature Photon.* **7**, 769–771 (2013).
3. Krishnamoorthy, K., Facchetti, A. F., Hu, W. & Bao, Z. *Phys. Chem. Chem. Phys.* **17**, 26509–26511 (2015).
4. Tang, C. G. *et al.* *Nature* **539**, 536–540 (2016).
5. Shen, Y., Hosseini, A. R., Wong, M. H. & Malliaras, G. G. *ChemPhysChem* **5**, 16–25 (2004).
6. Chueh, C.-C., Li, C.-Z. & Jen, A. K.-Y. *Energy Environ. Sci.* **8**, 1160–1189 (2015).
7. Zhou, N. *et al.* *Proc. Natl Acad. Sci. USA* **112**, 7897–7902 (2015).
8. Facchetti, A. *Chem. Mater.* **23**, 733–758 (2011).
9. Qiu, Y.-J. & Reynolds, J. R. *J. Electrochem. Soc.* **137**, 900–904 (1990).
10. Wu, Z. *et al.* *J. Am. Chem. Soc.* **138**, 2004–2013 (2016).

EVOLUTION

Insect invasions and natural selection

Observations of a real-time invasion of Australia by Asian honeybees demonstrate how natural selection can allow a small founding population to overcome the genetic odds stacked against success.

AMRO ZAYED

Social insects that live in colonies rank among the most notorious biological invaders. They have colonized almost all terrestrial ecosystems, threatening native faunas, domesticated livestock and human health¹. Yet the success of social-insect invaders is paradoxical, because they have a sex-determination system that gives rise to many sterile or inviable males in small founding

populations, which lack genetic diversity². Writing in *Nature Ecology & Evolution*, Gloag *et al.*³ show how, during the early stages of an invasion of the Asian honeybee *Apis cerana*, the action of natural selection can lessen this sex-determination problem, allowing even the smallest founding populations to thrive. The authors also provide a powerful example of natural selection observed in real time during biological invasions.

Sex determination in bees, like that in most

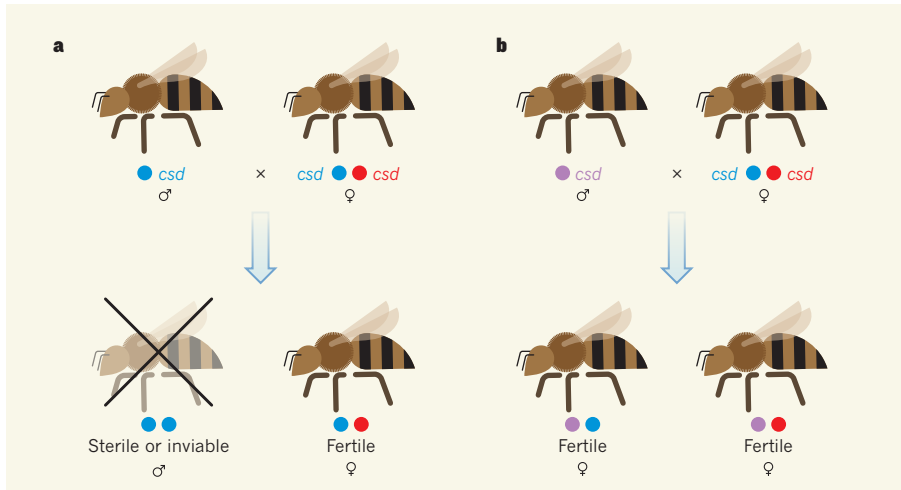


Figure 1 | Increased frequency of rare alleles by natural selection in an invasive bee population.

A female bee has two copies of the sex-determining *csd* gene and mates with a male bee, which has only one copy of the gene. The *csd* gene exists in different forms called alleles, shown in different colours.

a, If a female mates with a male that has an identical *csd* allele, half of the resulting offspring will have two identical *csd* alleles and will develop as sterile or inviable males that do not contribute to future generations. The offspring that have two different *csd* alleles develop as fertile females. **b**, All of the offspring from a mating between a male that has a rare allele, such as the one indicated in purple, and a female that does not have this rare allele will develop into fertile females. This would therefore increase the frequency of the rare allele in future generations, until all *csd* alleles have the same frequency. This type of natural selection — called balancing selection — was confirmed in observations made by Gloag *et al.*³ during a study of invading *Apis cerana* bees in Australia.

ants and wasps, is governed by the genetic make-up of a single, and usually highly variable, gene, called *csd* in honeybees⁴. Different versions of genes are called alleles. Haploid cells have one allele of each gene and diploid cells have two (these two alleles can be identical or different). In bees, unfertilized haploid eggs have a single *csd* allele and develop into haploid male bees. Fertilized, diploid eggs that have two different alleles of *csd* develop as female bees, which are workers or queens in honeybee colonies. However, under conditions in which there is limited genetic variation at *csd*, for example when the population is small, fertilized eggs can have two identical alleles of *csd*, thereby tricking the sex-determination machinery into producing a diploid male bee from a female-destined egg.

Diploid males are sterile or inviable in the vast majority of bee species studied so far². In honeybees, diploid males are quickly cannibalized by workers⁵. By converting a fraction of female-destined eggs into sterile or inviable males, the production of diploid males slows the growth of bee colonies that require female workers to function. In small, isolated populations, the production of diploid males can cause colony mortality⁶ and population extinction⁷. How do invasive social insects overcome this potential drawback of their unusual system of sex determination? To address this question, Gloag *et al.* tracked honeybee genetics during a nascent invasion of Australia by *A. cerana* over an eight-year period.

The authors speculated that a special form

of natural selection, called balancing selection, would have a role in reducing imbalances in the frequency of *csd* alleles. This would, in turn, reduce the production of inviable diploid males in the invasive population. With balancing selection, individuals that have rare *csd* alleles would be expected to have high fitness because they would be unlikely to mate with an individual that had the same allele — a mating that would result in the production of diploid male offspring (Fig. 1). Carriers of rare alleles should contribute more offspring to future generations and increase the frequency of their initially rare *csd* alleles over time. A prediction about balancing selection on *csd* is that an equilibrium would be reached when all *csd* alleles have the same frequency.

Gloag *et al.* painstakingly sequenced the functional parts of the *csd* gene in the *A. cerana* population to determine how it evolved during the invasion of Australia. They also analysed genetic diversity at several other regions of the bee genome. The invasion was first detected in late 2007, and the authors monitored the genetics of the invasive *A. cerana* population from 2008 to 2015. Ecological theory predicts that invasions initiated by large or multiple founding populations would be the most likely to succeed⁸. But Gloag *et al.* found that the genetic diversity in the invasive *A. cerana* population was consistent with a single introduction of a single colony.

In 2008, the authors identified only four different alleles at the *csd* gene, and they estimated that one out of four fertilized eggs would give rise to an inviable diploid male. However,

a year later, the authors documented three additional *csd* alleles. Gloag *et al.* propose that these three alleles were so rare that they had escaped detection in 2008. The authors showed that the rare alleles were not the result of new mutations because the alleles' DNA sequence was substantially different from those of the alleles found in 2008. It is also unlikely that these new *A. cerana* *csd* alleles were introduced by a secondary invasion, because the authors did not observe new alleles at the other DNA regions they examined. Rather, the presence of these three rare alleles in 2009 is best explained by balancing selection, in which bees that carried the rare *csd* alleles in 2008 contributed more offspring to the next generation.

Over the next six years, from 2009 to 2015, Gloag and colleagues found that the frequency of each of the seven *csd* alleles in the invasive population started to converge on an allele frequency of one-seventh, the theoretical frequency expected if balancing selection is assumed. By 2015, the authors estimated that one out of ten fertilized eggs developed into a diploid male, representing a 40% reduction in female mortality caused by diploid male production compared with the data from 2008 — this is a massive improvement in the fitness of the invasive population. Balancing selection at the *csd* gene has been indirectly inferred from studies of DNA-sequence change⁹ in native honeybee populations, and Gloag and colleagues' work provides a direct observation of this form of selection in real time.

Diploid males are typically found in invasive populations of ants¹⁰, bees¹¹ and wasps¹². The balancing selection at the *csd* gene documented by Gloag *et al.* might be a common process that enhances the success of social-insect invasions by correcting imbalances in the frequency of *csd* alleles that occur during founding events.

Although balancing selection clearly increased the fitness of invasive Asian honeybees in Australia, it is not clear whether this evolutionary force was essential for the successful establishment of the invasion. The invasive population of *A. cerana* still increased in size despite the skewed allele frequencies at the *csd* gene during the early stages of the invasion. Social-insect invaders might have other attributes that predispose them to be successful biological invaders. The intrinsic growth rates of some social insects might be so high that even a 25% increase in female mortality — the amount estimated by Gloag and colleagues to occur in the initial stages of *A. cerana*'s invasion — did not prevent population expansion.

Queen honeybees typically mate with 10–20 males, which spreads the consequences of diploid male production evenly across the invasive population; instead of some colonies producing a lot of diploid males, with others producing none, most colonies in the population would be expected to produce a

moderate level of diploid males¹³. Moreover, honeybee colonies recoup some of the costs of producing diploid males when they are cannibalized by workers. Perhaps the reduced genetic diversity found in invasive *A. cerana* reduces aggression between colonies as it does in invasive Argentine ants¹⁴. Nevertheless, Gloag and colleagues' study provides a clear example of how rapid evolutionary changes can affect the fitness of invasive populations. ■

Amro Zayed is in the Department of Biology, York University, Toronto, Ontario M3J 1P3, Canada.
e-mail: zayed@yorku.ca

- Holway, D. A., Lach, L., Suarez, A. V., Tsutsui, N. D. & Case, T. J. *Annu. Rev. Ecol. Syst.* **33**, 181–233 (2002).
- Harpur, B. A., Sobhani, M. & Zayed, A. *Entomol. Exp. Appl.* **146**, 156–164 (2013).

- Gloag, R. *et al.* *Nature Ecol. Evol.* <http://dx.doi.org/10.1038/S41559-016-0011> (2016).
- Beye, M., Hasselmann, M., Fondrk, M. K., Page, R. E. Jr & Omholt, S. W. *Cell* **114**, 419–429 (2003).
- Woyke, J. J. *Apicult. Res.* **18**, 122–127 (1979).
- Plowright, R. C. & Pallett, M. J. *Can. Entomol.* **111**, 289–294 (1979).
- Zayed, A. & Packer, L. *Proc. Natl Acad. Sci. USA* **102**, 10742–10746 (2005).
- Simberloff, D. *Annu. Rev. Ecol. Syst.* **40**, 81–102 (2009).
- Hasselmann, M. & Beye, M. *Proc. Natl Acad. Sci. USA* **101**, 4888–4893 (2004).
- Ross, K. G., Vargo, E. L., Keller, L. & Trager, J. C. *Genetics* **135**, 843–854 (1993).
- Zayed, A., Constantini, S. A. & Packer, L. *PLoS ONE* **e868** (2007).
- Liebert, A. E., Sumana, A. & Starks, P. T. *Biol. Lett.* **1**, 200–203 (2005).
- Cook, J. M. & Crozier, R. H. *Trends Ecol. Evol.* **10**, 281–286 (1995).
- Holway, D. A., Suarez, A. V. & Case, T. J. *Science* **282**, 949–952 (1998).

This article was published online on 7 November 2016.

NANOSCIENCE

Flexible graphene strengthens friction

Previous observations showed that friction on graphene increases gradually when a probe starts to slide across the material's surface. Simulations now reveal that this effect is related to bending of the graphene sheet. [SEE LETTER P.541](#)

ASTRID S. DE WIJN

As anyone who has slipped on wet autumn leaves might realize, friction in layered materials can behave in peculiar ways. Such friction can be studied on small scales by looking at graphene, a material that consists of single sheets of carbon atoms. Experiments^{1,2} performed a few years ago using atomic force microscopy (AFM) revealed that friction is larger when the number of stacked graphene sheets is smaller. Moreover, when a probe begins to move across the surface of stacked graphene sheets, friction initially increases and then levels off². A wide range of suggestions has been proposed to explain these observations, but on page 541, Li *et al.*³ solve the mystery using computer simulations. They report that graphene's strange behaviour is related to distortions that strengthen friction on thin stacks of sheets.

Friction is so common in our daily lives that we barely think about it. Humans have been forced to deal with it since prehistoric times: the wheel is essentially a device for reducing friction when moving heavy loads, and rubbing sticks together to make fire makes use of the fact that friction generates heat. We have therefore developed a massive amount of phenomenological knowledge of friction under

specific circumstances, and especially of how to reduce it. But our fundamental understanding of how and why friction works is still sorely lacking.

Layered materials are a prime example of this knowledge gap. Some of them, such as graphite — stacked layers of graphene — and molybdenum disulfide are commonly used in powdered form as solid lubricants or as additives in liquid lubricants, and sometimes as low-friction coatings. We use these materials all the time, but are only now starting to understand their workings.

Graphene has received much attention in the past few years (see ref. 4, for example), largely because of its remarkable electronic properties — research into this material was rewarded with the Nobel Prize in Physics in 2010. But it is graphene's mechanical properties that are interesting when considering its friction. Most importantly for Li and colleagues' results, single sheets of layered materials bend easily. Although it might seem obvious that such bending must have a role in the friction of thin stacks of graphene, many explanations for the role of bending were possible.

To understand Li and co-workers' results, we need to know what surfaces in contact look like. Real surfaces are rough on microscopic scales, even if they look flat to the naked eye.

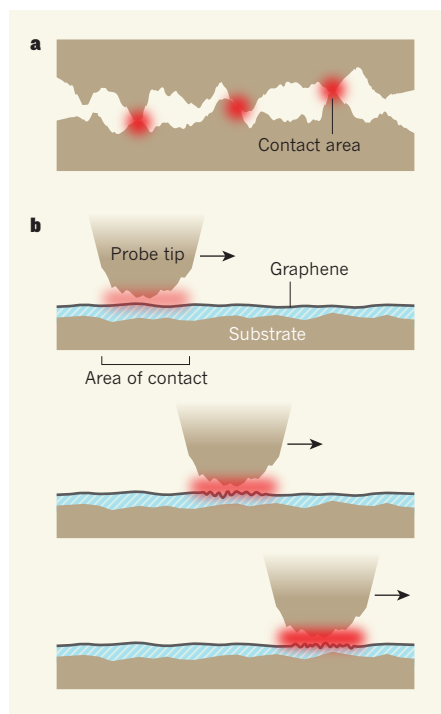


Figure 1 | Friction on graphene sheets. **a**, Surfaces that look smooth on a macroscopic scale are rough on the microscopic scale. Contact between surfaces thus occurs only between microscopic projections from the surfaces, and friction between the surfaces depends on the strength of the interactions between the contacting projections. **b**, Li *et al.*³ performed simulations to investigate what happens when a nanometre-scale probe is dragged along the surface of graphene on a silicon substrate. During the initial stages of sliding, the strength of the probe's interaction with the graphene increases gradually, thus increasing the friction between the probe and the graphene. The substrate-graphene interaction (blue) also plays a part in the observed effects, although its exact role is unclear.

Two surfaces in contact therefore never fit perfectly against each other, but meet only where microscopic projections come into contact (Fig. 1a). The actual area in contact is thus typically many orders of magnitude smaller than the apparent contact area.

These microscopic projections play a central part in friction. More specifically, the strength of the interaction between the microscopic contacts, which Li *et al.* refer to as the quality of the contact, is a key factor in the dissipation of the kinetic energy of sliding. The interaction strength can be affected by many factors, such as a material's properties, its chemistry and a mismatch of the crystal lattices in the contacting surfaces⁵, but also, in the case of graphene, by bending⁶.

The previously reported AFM experiments^{1,2} measured what happens at a single contact on stacks of graphene. Li *et al.* simulated those experiments numerically and investigated the strength of the interaction at a single contact. In the simulations, they put

one or several layers of graphene on top of a substrate (silicon) and rubbed the surface of the material with a tiny probe. They observed that bending distortions of the graphene sheets under the probe gradually increase as the probe moves; this increases the contact strength, and thereby also the friction (Fig. 1b). The effect is bigger if there are fewer layers, because thin stacks of graphene are more flexible than thicker stacks. The authors also find that the substrate underneath the graphene contributes to the increased contact strength.

It is worth noting that the authors checked carefully to make sure that the effect observed in their simulations was not related to the sliding speed. This was crucial, because current limitations to computational power make it impossible to simulate sliding at the low speeds used in AFM experiments.

Despite the fresh insight provided by Li *et al.*, many questions remain. What exactly causes the bending distortions to grow? And what is the role of the substrate under the graphene? When the authors altered the parameters of their model to increase the binding of the substrate to graphene, the contact strength no longer changed as the probe moved. The same was true when there was no substrate at all.

Moreover, both the substrate and the probe used by the authors were amorphous materials that lacked any structural order. It would

be interesting to see what happens if the simulations are repeated using an ordered substrate and/or probe. Graphene is often studied on ordered substrates in other contexts, and such a study might also provide insight into the mechanisms at play if the structures of the substrate and tip affect strengthening in some way.

Friction is a messy problem to study, because its effects are often found in highly complex systems that involve a variety of materials and chemistry. It is usually not clear which of these are essential for a specific friction effect, and which have little or no role. This makes meaningful study of such phenomena difficult. Li and colleagues' finding of increases in contact strength in a relatively simple system opens up the possibility of studying the strengthening process in a much more controlled way. ■

Astrid S. de Wijn is in the Department of Engineering Design and Materials, Norwegian University of Science and Technology, 7491 Trondheim, Norway.
e-mail: astrid.dewijn@ntnu.no

1. Filleter, T. *et al.* *Phys. Rev. Lett.* **102**, 086102 (2009).
2. Lee, C. *et al.* *Science* **328**, 76–80 (2010).
3. Li, S. *et al.* *Nature* **539**, 541–545 (2016).
4. Editorial. *Nature Mater.* **10**, 1 (2011).
5. Dienwiebel, M. *et al.* *Phys. Rev. Lett.*, **92**, 126101 (2004).
6. van Wijk, M. M., Dienwiebel, M., Frenken, J. W. M. & Fasolino, A. *Phys. Rev. B* **88**, 235423 (2013).

CELL BIOLOGY

A mitochondrial brake on vascular repair

Injured blood vessels are repaired by vascular smooth-muscle cells. It emerges that the protein Fat1 regulates the proliferation of these cells by inhibiting the function of mitochondria. SEE LETTER P.575

CHARLES E. DE BOCK & RICK F. THORNE

After injury to the cells that line blood-vessel walls, vascular smooth-muscle cells (VSMCs) move to the injured region and proliferate to cover the damaged site. However, cellular overproliferation at the repair site might cause vessel-wall thickening that reduces blood flow through the vessel. In addition to naturally occurring damage, blood-vessel injury can be a consequence of surgical interventions such as transplants or procedures to open narrowed blood vessels. An understanding of the signals that regulate VSMC proliferation might enable the development of clinical approaches to limit this process and prevent blood-vessel narrowing. Cao *et al.*¹ demonstrate on page 575 that the Fat1 receptor protein, a negative regulator of VSMC

proliferation, has a direct and unexpected role in regulating energy production in mitochondria, the organelles that act as cellular powerhouses.

Fat1 is a member of the vertebrate Fat cadherins, a small family of proteins whose primary function is unclear². Previous work³ revealed that Fat1 enhances the migration and limits the proliferation of VSMCs, providing clues that Fat1 could have a role in vascular-cell remodelling.

The Fat1 receptor is normally present at the cell surface, and its structure consists of an extracellular domain, a transmembrane segment and an intracellular domain². Cao *et al.* used mass spectrometry to identify proteins that interact with the intracellular domain of mouse Fat1. They discovered that 22 Fat1 interactors are proteins of



50 Years Ago

Recent reports on the occurrence of “pink spot” in the urines of schizophrenics are somewhat conflicting and cast doubt on its chemical identity and its relationship to schizophrenia ... The possibility that “pink spot” may be associated with brain damage induced by drugs is supported by the fact that it is also found in high concentration in the urines of patients suffering from Parkinsonism, a disease in which the brain is indisputably damaged. This leaves open the questions of the relationship of “pink spot” to schizophrenia and also its relationship to dopamine metabolism ... it would be interesting to know whether any of the patients who have been reported in the literature as “pink spot” excretors also showed clinical signs of brain damage in the form of dyskinesia induced by drugs at the time of investigation or whether they have developed it since the reports were published.

From *Nature* 26 November 1966

100 Years Ago

Whatever difference of opinion may exist in regard to the stimulus which has moved the Government to take control of our food supplies, all are agreed that it has not come about a moment too soon ... The provisions outlined will empower the Food Controller to take measures both preventive and regulatory ... But Government measures cannot stop with the mere regulation of food supplies. Powers must be taken to compel a greater production of home food and to ensure a larger acreage of wheat ... Objection may be raised to the shortage of labour ... why should prisoners of war not be utilised to the fullest degree possible in raising the home production of food?

From *Nature* 23 November 1916

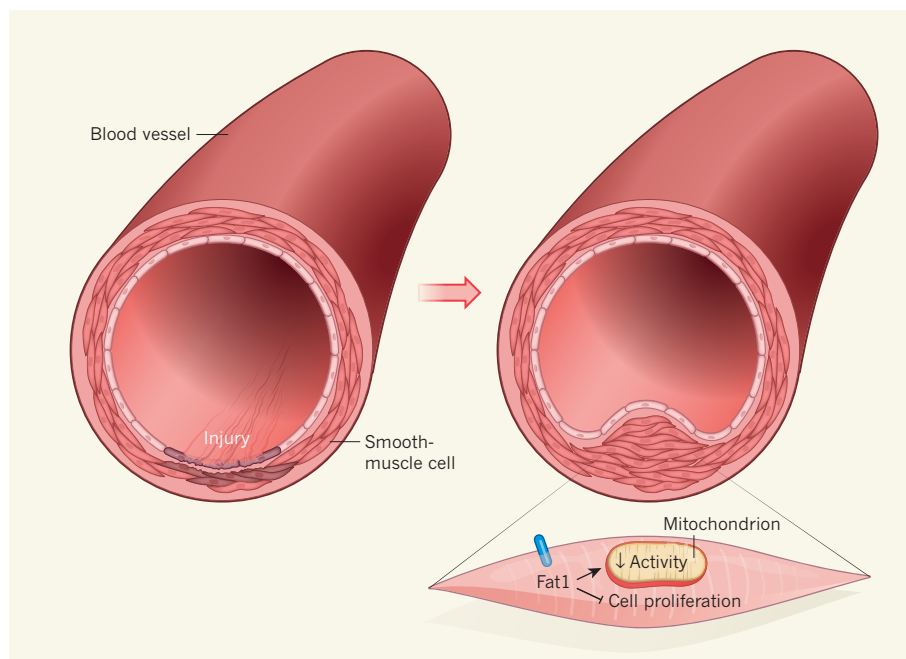


Figure 1 | Fat1 goes with the flow. Injury to the walls of blood vessels can occur through natural causes or surgical interventions. Vascular smooth-muscle cells (VSMCs) migrate to the damaged site and proliferate as part of the healing response. However, this can result in cellular overproliferation, causing vessel-wall thickening and narrowing the blood vessel. The Fat1 receptor protein (blue) limits the proliferation of VSMCs³. In studies using mice and human cells, Cao *et al.*¹ report that a portion of Fat1 interacts with proteins in the mitochondria and inhibits mitochondrial activity in VSMCs.

the inner mitochondrial membrane.

Many biochemists would be cautious about further investigating any such interactions with mitochondrial proteins because non-specific interactions with these proteins are common. Nonetheless, Cao *et al.* tested and validated their findings using a variety of techniques. The authors observed that Fat1 is present inside the mitochondria of VSMCs. They also found that non-full-length portions of Fat1 are enriched in mitochondria, and that these portions interact directly with proteins of the mitochondrial respiratory machinery that generate the energy-carrying nucleotide ATP. Subsequent analyses revealed that mouse VSMCs lacking Fat1 had enhanced mitochondrial respiration in the absence of other alterations to mitochondrial structure or mass that might be responsible for such enhancement. Cao and colleagues conclude that Fat1 dampens mitochondrial respiration to suppress VSMC proliferation (Fig. 1).

The importance of mitochondrial function to the proliferative capacity of VSMCs was already known⁴. But how central is Fat1 signalling to the mitochondrial regulation needed to enable VSMC repair of vascular damage? Cao *et al.* established the relevance of this pathway in human tissue through *in vitro* studies demonstrating that human FAT1 can regulate mitochondrial respiration and proliferation in human VSMCs from arterial blood vessels. The authors also observed FAT1 expression in human VSMCs in an artery-repair context, when they tested post-mortem artery samples

from people who had undergone surgical insertion of a stent device to correct atherosclerosis, an artery-narrowing condition.

Cao and colleagues investigated a model of arterial injury using mice whose smooth-muscle cells lacked the *Fat1* gene. Injured arteries in these mice became significantly narrower than did those of control mice, establishing that the presence of Fat1 normally limits the potential of VSMCs to contribute to vascular blockage. However, although this model system might be relevant to many aspects of vascular-cell proliferation and remodelling, it does not reflect all the processes involved in vascular disease. Nevertheless, this work establishes a framework with which to further investigate the Fat1 and mitochondrial signalling pathway — for example, by testing whether this mitochondrial-control pathway is evolutionarily conserved in other members of the Fat protein family.

Vascular biology is not the only relevant setting for Fat1 investigation. Roles for Fat1 have been proposed in a surprisingly diverse range of disease states^{1,2}. Cancer is of particular interest, given that studies⁵ in the fruit fly *Drosophila melanogaster* indicate that *Ft*, a fly version of *Fat1*, is a tumour-suppressor gene. A tumour-suppressor function is consistent with reports of repression or inactivation of FAT1 in certain human cancers^{2,6} — for example, loss of FAT1 can result in tumour-promoting Wnt-protein signalling⁶.

However, FAT1 is overexpressed in some types of cancer⁷, creating a conundrum about

whether this protein promotes or inhibits tumours. FAT1 might be overexpressed in cancer because tumour cells need to escape destruction by cell-death processes, and high levels of FAT1 protect against cell death⁸. FAT1 is also upregulated in low-oxygen (hypoxic) conditions⁹. The stressful cellular environment that results from hypoxia often occurs in solid tumours, and perhaps FAT1 upregulation allows tumour cells to survive in such conditions. As is so often the case in biology, the determination of a protein's action requires observations to be placed in the correct cellular context. The challenge is to discover whether Fat1's function in controlling mitochondrial respiration is relevant to human cancer.

Many questions remain. Previous work¹⁰ shows that *D. melanogaster* Ft enhances mitochondrial function, which contrasts with Cao and colleagues' observations that Fat1 inhibits mitochondrial function in the mammalian systems they studied. Whether this difference results from evolutionary divergence or reflects a cell-dependent context is unknown.

Discovering how Fat1 is regulated is a

key area for further investigation. Cao *et al.* observed that smaller portions of Fat1, rather than the full-length protein, are associated with the mitochondria. This observation suggests that Fat1 undergoes enzyme-mediated cleavage to enter the mitochondria. Fat1 is a substrate of the γ -secretase protein complex¹¹ — the protease enzyme that cleaves and releases cytoplasmic domains of transmembrane proteins¹² — although the authors did not investigate Fat1 cleavage by this or any other enzyme.

It is also not known whether a ligand molecule binds the extracellular domain of Fat1. Given that endothelial cells in blood-vessel walls sense mechanical damage and relay this information to adjacent VSMCs¹³, endothelial cells might signal to Fat1 by an unknown ligand. An understanding of such aspects of FAT1 regulation will probably be needed for the therapeutic potential of this signalling pathway to be exploited in the clinic. ■

Charles E. de Bock is at the VIB Center for the Biology of Disease, Leuven 3000, Belgium,

and KU Leuven Center for Human Genetics, Belgium. Rick F. Thorne is at the Hunter Medical Research Institute, New Lambton Heights, New South Wales 2305, Australia, and the University of Newcastle, Australia.

e-mail: rick.thorne@newcastle.edu.au

1. Cao, L. L. *et al.* *Nature* **539**, 575–578 (2016).
2. Sadeqzadeh, E., de Bock, C. E. & Thorne, R. F. *Med. Res. Rev.* **34**, 190–221 (2014).
3. Hou, R., Liu, L., Anees, S., Hiroyasu, S. & Sibinga, N. E. S. *J. Cell Biol.* **173**, 417–429 (2006).
4. Chiong, M. *et al.* *Front. Cell Dev. Biol.* **2**, 72 (2014).
5. Mahoney, P. A. *et al.* *Cell* **67**, 853–868 (1991).
6. Morris, L. G. T. *et al.* *Nature Genet.* **45**, 253–261 (2013).
7. de Bock, C. E. *et al.* *Leukemia* **26**, 918–926 (2012).
8. Kranz, D. & Boutros, M. *EMBO J.* **33**, 181–197 (2014).
9. Madan, E. *et al.* *Int. J. Cancer* **139**, 2570–2582 (2016).
10. Sing, A. *et al.* *Cell* **158**, 1293–1308 (2014).
11. Magg, T., Schreiner, D., Solis, G. P., Bade, E. G. & Hofer, H. W. *Exp. Cell Res.* **307**, 100–108 (2005).
12. Bray, S. J. *Nature Rev. Mol. Cell Biol.* **17**, 722–735 (2016).
13. Gao, Y., Chen, T. & Raj, J. U. *Am. J. Respir. Cell Mol. Biol.* **54**, 451–460 (2016).

This article was published online on 9 November 2016.

MATERIALS SCIENCE

Improving the image of nanoparticles

A biocompatible probe that combines fluorescent nanodiamonds and gold nanoparticles allows cells to be imaged using both optical and electron microscopy techniques, opening up fresh opportunities for biological research.

CHRISTOPHER S. WOOD
& MOLLY M. STEVENS

Diamonds are renowned for their sparkle, but over the past few years many scientists have become interested in nanometre-scale fragments of diamonds that fluoresce. Such nanodiamonds are attractive probes for fluorescent cell imaging¹ because they fluoresce intensely, show low toxicity to cells and are stable to chemicals and light. However, they cannot easily be tracked using electron microscopy, limiting their utility as biological probes. Writing in *Nano Letters*, Liu *et al.*² report that they have combined fluorescent nanodiamonds with gold nanoparticles to create 'all-in-one' hybrid particles that can be used in both fluorescent imaging and electron microscopy. This potentially opens the way to studying cellular processes such as endocytosis — a membrane-dependent process that delivers materials to cells — in a single sample at multiple resolutions.

The use of nanoparticle-based probes is the latest in a long line of developments in imaging

technology. Although electron microscopy can resolve structures smaller than is achievable using light, the samples studied must be fixed or frozen, providing limited insight into the dynamic processes of living systems. Super-resolution microscopy techniques³ based on fluorescent imaging overcome this limitation and allow the visualization of biological processes at work, even at the molecular level^{4,5}. However, these techniques are limited by the brightness of the required fluorophores (fluorescent proteins or small molecules), and by the light and chemical stability of those molecules.

Researchers are increasingly turning to nanoparticle-based imaging probes to overcome these obstacles. Matter on the nanoscale can give rise to phenomena such as intense fluorescence, which minimizes some of the drawbacks associated with molecular fluorophores. For example, fluorescent nanodiamonds (FNDs) contain lattice defects known as single nitrogen-vacancy centres, which fluoresce at wavelengths of about 700 nanometres when excited by a laser.

Nanoparticles also provide additional modes of imaging and methods for probing their local environments — such as surface-enhanced Raman scattering⁶ or nanoscale magnetometry⁷ (the detection of magnetic fields at nanoscale resolution). However, these abilities are rarely combined in a single particle.

Gold nanoparticles (AuNPs) are extensively used in biosensing and bioimaging because they enhance the contrast of structures in electron microscopy. Many sizes of AuNPs quench the excited states of fluorescent molecules, an effect that has been exploited in numerous highly sensitive biosensors⁸. This otherwise useful property limits our ability to track AuNPs in living systems using fluorescence microscopy. But in Liu and co-workers' hybrid nanoparticles, FNDs fluoresce even though they are in close proximity to similarly sized AuNP contrast agents. AuNPs and FNDs have been combined before⁹, but the current work demonstrates the potential of hybrid-particle constructs to provide multiple modes of imaging or sensing at a single point in a sample.

Developing complex biocompatible probes and biosensors from multiple nanoscale components is challenging. The difficulty arises partly because nanoparticles tend to form aggregates, especially in biological environments, but also because it is hard to form and purify stable particle constructs that contain a defined number of components. To overcome these problems, Liu *et al.* associated the AuNPs with the FNDs by mixing them with the protein human serum albumin (which interacts with FNDs and AuNPs), and stabilizing the resulting complexes by forming covalent crosslinks between the AuNPs and the protein and between the protein and the

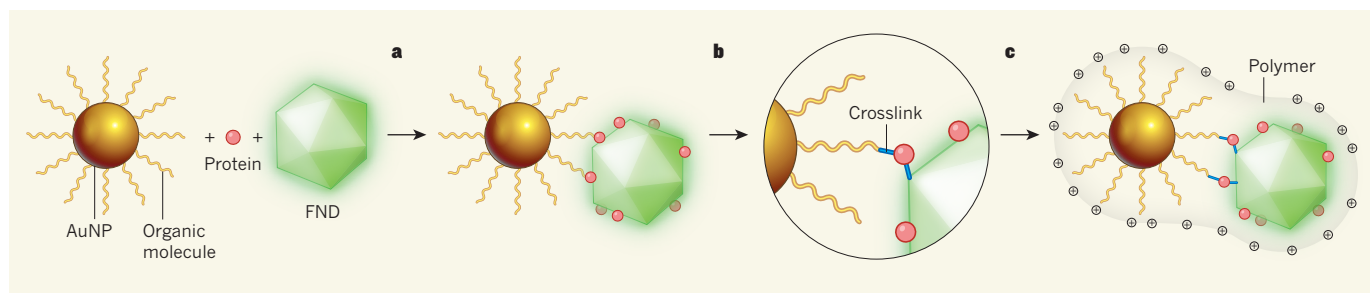


Figure 1 | Preparation of hybrid nanoparticles for biological imaging studies. Liu *et al.*² have made nanoparticles that contain fluorescent nanodiamonds (FNDs) and gold nanoparticles (AuNPs). **a**, They began by mixing AuNPs (decorated with organic molecules to prevent their aggregation) with FNDs and a protein (human serum albumin). This generated hybrid particles composed of an FND with protein molecules adsorbed to its surface, along with one or two AuNPs. **b**, The authors then

stabilized the hybrid particles by forming covalent crosslinks between the organic groups on the AuNPs and the proteins, and between the proteins and the FNDs. **c**, Finally, the hybrid particles were coated with a positively charged polymer, enabling them to be taken up efficiently by cells. The FND component allows the particles to be tracked in cells using fluorescence imaging, whereas the AuNP components enable the particles to be imaged using electron microscopy. (Adapted from ref. 2.)

FNDs (Fig. 1). The authors then purified the desired products by ultracentrifugation, and coated them with a positively charged polymer to further stabilize them and to enhance cell uptake. The resulting particles contained one or two AuNPs per FND, were readily absorbed by cells and showed low cytotoxicity.

Intriguingly, covalently linking the AuNPs to the FND led to an improvement in FND fluorescence and a reduction in the lifetime of excited states in the FNDs, both of which occur as a result of coupling between the nitrogen-vacancy centres and AuNP surface plasmons (collective oscillations of delocalized electrons). The authors also attribute the increased fluorescence to concentration of the exciting laser pulses by the AuNPs, leading to more-efficient laser excitation of the hybrid particles — although they do not discuss a precise mechanism for this.

The enhanced fluorescence is particularly interesting because it might improve the signal-to-noise ratio of FNDs in sensing and imaging based on nanoscale magnetometry⁷. The nitrogen-vacancy centres found in FNDs are particularly sensitive to external magnetic fields, and so nanoscale magnetometry using FNDs is emerging as a potentially useful sensing technique in biological environments¹⁰. It could enable the detection of single electrons, for example, and thus provide insights into biological processes, such as respiration, that involve electron transfer or the recombination of free radicals¹¹.

The development of Liu and colleagues' biocompatible hybrid-nanoparticle probes is an important step towards improved imaging of cell mechanisms, and opens up fresh possibilities for sensing within cells. For example, the authors observed both the uptake of FND–AuNPs by cells using optical microscopy, and the release of the nanoparticles from cellular vesicles called endosomes into the cytosol using transmission electron microscopy. However, it remains to be seen whether the hybrid nanoparticles can be used *in vivo*. It is also unclear whether the authors' approach

for attaching AuNPs to FNDs can be used to attach other nanometre-sized structures, such as antibodies, enzymes and further types of nanoparticle.

A robust, modular method for synthesizing hybrid nanoparticles in general — not just those involving FNDs — is now needed to allow scientists to rapidly combine different imaging probes and biosensors without extensive optimization. This would enable biological environments to be probed in many ways at specific points within a sample. Such a method would also enable nanoparticle-based drug-delivery systems to be incorporated easily into hybrid probes. This could lead to rapid prototyping of 'theranostic' agents, in which particles deliver therapeutic compounds to patients and report the effects using several forms of information. ■

Christopher S. Wood and Molly M. Stevens
are in the Departments of Materials and

Bioengineering, and at the Institute for Biomedical Engineering, Imperial College London, London SW7 2AZ, UK.
e-mail: m.stevens@imperial.ac.uk

1. Yu, S.-J., Kang, M.-W., Chang, H.-C., Chen, K.-M. & Yu, Y.-C. *J. Am. Chem. Soc.* **127**, 17604–17605 (2005).
2. Liu, W. *et al. Nano Lett.* **16**, 6236–6244 (2016).
3. Fernández-Suárez, M. & Ting, A. Y. *Nature Rev. Mol. Cell Biol.* **9**, 929–943 (2008).
4. Wegel, E. *et al. Sci. Rep.* **6**, 27290 (2016).
5. Henriques, R., Griffiths, C., Rego, E. H. & Mhlanga, M. M. *Biopolymers* **95**, 322–331 (2011).
6. Howes, P. D., Rana, S. & Stevens, M. M. *Chem. Soc. Rev.* **43**, 3835–3853 (2014).
7. Balasubramanian, G. *et al. Nature* **455**, 648–651 (2008).
8. Howes, P. D., Chandrawati, R. & Stevens, M. M. *Science* **346**, 1247390 (2014).
9. Liu, Y. L. & Sun, K. W. *Appl. Phys. Lett.* **98**, 153702 (2011).
10. McGuinness, L. P. *et al. Nature Nanotechnol.* **6**, 358–363 (2011).
11. Wu, Y., Jelezko, F., Plenio, M. B. & Weil, T. *Angew. Chem. Int. Edn* **55**, 6586–6598 (2016).

This article was published online on 16 November 2016.

PLANT SCIENCE

A war over water when bacteria invade leaves

Plants and bacteria battle for control of water during leaf infection, as is demonstrated by a bacterial species that manipulates plant cells to create a water-rich environment that promotes bacterial growth. [SEE ARTICLE P. 524](#)

GWYN A. BEATTIE

Invasive bacterial leaf pathogens disarm a plant's defence mechanisms by injecting what are known as effector proteins directly into plant cells. In addition to the plant's defence mechanisms, invading bacteria might encounter another obstacle — a severe shortage of water in the leaf intercellular space

(apoplast) that limits bacterial growth. On page 524, Xin *et al.*¹ report that *Pseudomonas syringae* bacteria inject effector proteins that specifically target plant-cell processes to promote apoplast hydration, thereby creating an environment suitable for bacterial growth.

High humidity favours the development of bacterial leaf diseases that are associated with the formation of leaf spots. An early symptom

of leaf-spot diseases is water soaking², a localized abundance of water in the apoplast. Extensive water soaking of plants is associated with epidemics and severe disease symptoms³.

To investigate the relationship between the water soaking of leaves and bacterial growth, Xin and colleagues used microscopy to analyse the thale cress plant *Arabidopsis thaliana*. They observed that water soaking occurred only transiently during the start of leaf-spot development. The leaf sites that exhibited this transient water soaking were also the sites of early bacterial replication, and ultimately became the locations of disease lesions. These results provide direct visual evidence that links water soaking to bacterial growth and leaf-spot development.

The WstE and AvrE effector-protein family can promote water soaking in some plant species^{4,5}, although the mechanism involved and the role of water soaking in plant infection had not been identified for these proteins. Xin *et al.* found that the *P. syringae* effector proteins AvrE and HopM1 can each promote water soaking in *A. thaliana*, and that this soaking occurs regardless of whether the proteins are expressed by the bacteria or expressed in the plant cell by genetic engineering. The authors demonstrated that HopM1 induces soaking through a pathway that involves degradation of the plant protein MIN7. This discovery indicates that a bacterial pathogen can actively promote hydration of the apoplast by modifying a plant-cell target.

How water accumulates in the apoplast after MIN7 is destroyed is not known. Possible mechanisms include changes in cell-membrane function and stability, changes to the aquaporins — water-transport proteins — in cell membranes, or alterations to ion-transport activity across the membrane. A role for the cell membrane as a target for pathogen-mediated effects on water movement is supported by the facts that MIN7 is involved in maintaining a normal cell membrane⁶ and that AvrE moves to the cell membrane after injection⁷.

Xin *et al.* found that only two plant functions must be altered to enable bacterial infection: a plant immune-defence system known as pattern-triggered immunity and water limitation in the apoplast. A plant mutant lacking both MIN7 and pattern-triggered immunity enabled strong growth, not only of a non-pathogenic *P. syringae* strain, but also of other bacterial communities normally resident at low levels on leaves. The growth of these resident communities in the apoplast correlated with the appearance of leaf yellowing and small regions of plant-cell death, revealing that even bacteria not known to be pathogenic can exact a detectable toll on plant health if the conditions arise to enable excessive bacterial proliferation.

The humidity in the air outside the leaf can influence a plant's ability to maintain a water-limited apoplast, because this ambient

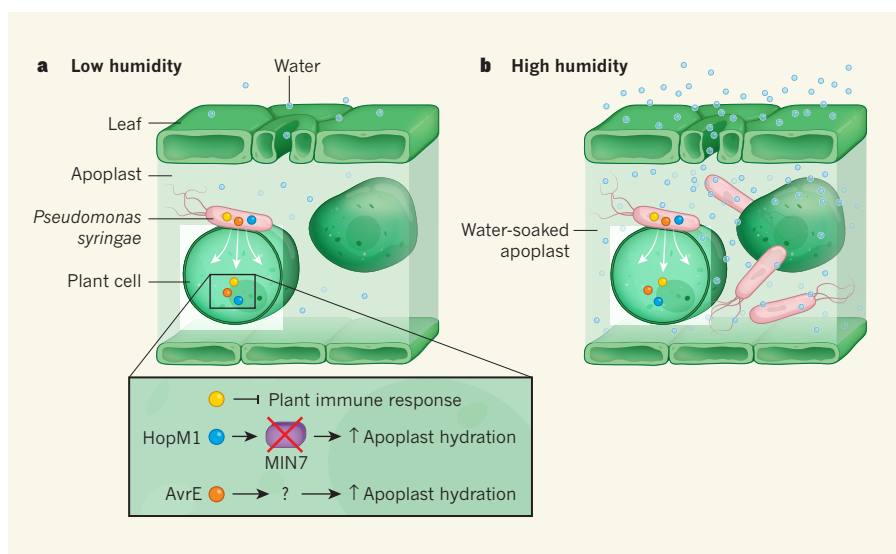


Figure 1 | A bacterial pathogen can affect the hydration around leaf cells. **a**, Bacteria need water to grow, and plants minimize bacterial growth in leaves by limiting water in the space between leaf cells (the apoplast). Pathogenic bacteria such as *Pseudomonas syringae* inject a variety of proteins directly into plant cells to aid invasion. These proteins include inhibitors of the plant immune system. Xin *et al.*¹ report that two injected bacterial proteins, HopM1 and AvrE, can each induce apoplast hydration. The authors demonstrate that HopM1 degrades MIN7, a plant protein that has a key role in limiting fluid loss from plant cells. However, under conditions of low ambient humidity, the bacterial injection of HopM1 and AvrE may induce only low levels of apoplast hydration. The mechanism by which AvrE induces apoplast hydration is unknown, but might involve targeting of a host protein complex¹⁰. **b**, Under conditions of high ambient humidity, HopM1 and AvrE induce high levels of apoplast hydration known as water soaking. Water soaking in the apoplast is essential for bacterial proliferation and for the development of symptoms of plant disease.

air connects directly to the apoplast through pores called stomata. Low-humidity air should help to limit the water in the apoplast and aid in antibacterial defence. This is consistent with the low levels of leaf-spot disease observed in arid conditions. By contrast, continuous water soaking should negate a defence strategy of maintaining low apoplast hydration, and this is consistent with the abundant growth of non-pathogenic bacteria in the apoplast of persistently water-soaked leaves⁵.

Xin and colleagues tested different ambient humidity conditions, and demonstrated that the humidity must be high to enable the pathogen to induce water soaking that is dependent on the injection of AvrE or HopM1 (Fig. 1). This strong requirement for humidity reflects a fundamental tenet of plant pathology: disease depends on the host, the pathogen and the environment.

Plants and pathogens battle for water during the initial pathogenic attack, and also later, after the plants have been infected. Xin and colleagues found that MIN7 is stabilized by a type of plant immune response known as effector-triggered immunity. This indicates that the plants have evolved mechanisms to counteract the bacterially induced destruction of MIN7 that enables apoplast hydration. The ability of plants to affect such hydration is consistent with the low apoplast hydration observed in *A. thaliana* plants that can resist an attack by *P. syringae*⁸. Reduced fluid movement into the apoplast during effector-triggered

resistance of *A. thaliana* to *P. syringae* attack suggests that, as well as preventing apoplast hydration, plants might also actively promote drying of the apoplast⁹.

The authors' study supports a model in which plants maintain a water-limited apoplast as a barrier against bacterial growth, and in which pathogen effectors disarm this barrier, leading to water soaking in conditions of high humidity. A key next step in understanding this fascinating plant–pathogen battle for control of water will be to identify the nature of the 'flood gates' to the apoplast and the molecular mechanisms by which pathogens open them. ■

Gwyn A. Beattie is in the Department of Plant Pathology and Microbiology, Iowa State University, Ames, Iowa 50011, USA.
e-mail: gbeattie@iastate.edu

1. Xin, X. F. *et al.* *Nature* **539**, 524–529 (2016).
2. Davis, K. R., Schott, E. & Ausubel, F. M. *Mol. Plant-Microb. Interact.* **4**, 477–488 (1991).
3. Beattie, G. A. *Annu. Rev. Phytopathol.* **49**, 533–555 (2011).
4. Asselin, J. E. *et al.* *Plant Physiol.* **167**, 1117–1135 (2015).
5. Ham, J. H., Majerczak, D. R., Arroyo-Rodriguez, A. S., Mackey, D. M. & Coplin, D. L. *Mol. Plant-Microb. Interact.* **19**, 1092–1102 (2006).
6. Tanaka, H., Kitakura, S., De Rycke, R., De Groodt, R. & Friml, J. *Curr. Biol.* **19**, 391–397 (2009).
7. Xin, X.-F. *et al.* *Plant Physiol.* **169**, 793–802 (2015).
8. Wright, C. A. & Beattie, G. A. *Proc. Natl Acad. Sci. USA* **101**, 3269–3274 (2004).
9. Freeman, B. C. & Beattie, G. A. *Mol. Plant-Microb. Interact.* **22**, 857–867 (2009).
10. Jin, L. *et al.* *PLoS Pathog.* **12**, e1005609 (2016).

Emergent phenomena induced by spin–orbit coupling at surfaces and interfaces

Anjan Soumyanarayanan^{1,2}, Nicolas Reyren³, Albert Fert³ & Christos Panagopoulos¹

Spin–orbit coupling (SOC) describes the relativistic interaction between the spin and momentum degrees of freedom of electrons, and is central to the rich phenomena observed in condensed matter systems. In recent years, new phases of matter have emerged from the interplay between SOC and low dimensionality, such as chiral spin textures and spin-polarized surface and interface states. These low-dimensional SOC-based realizations are typically robust and can be exploited at room temperature. Here we discuss SOC as a means of producing such fundamentally new physical phenomena in thin films and heterostructures. We put into context the technological promise of these material classes for developing spin-based device applications at room temperature.

The electric field experienced by a travelling electron translates, in its rest frame, to a magnetic field proportional to its velocity—a relativistic effect which is notable in crystalline lattices with heavy atoms. The Zeeman interaction between the electron spin and this effective magnetic field is equivalent to the coupling of the electronic spin and momentum degrees of freedom, known as SOC. SOC can split degenerate bands with finite angular momentum (p , d and f), modifying the electronic band structure. Importantly, SOC effects are greatly enhanced in reduced dimensions (Fig. 1, left and right). First, inversion symmetry is broken at the surface or interface, and the resultant electric field couples to the spin of itinerant electrons. This phenomenon, known as Rashba SOC¹, produces spin-split dispersion even at the surfaces of conventional metals (such as Au and Bi)². Recently discovered topological insulators, have spin-polarized surface states with additional topological properties. In both these cases, strong two-dimensional (2D) SOC locks the electron spin and momentum.

Spin–momentum locking in 2D geometries has direct consequences for the interplay between the charge and spin transport (Fig. 1, top left). An in-plane charge current induces a transverse spin accumulation (uniform non-zero spin density). This spin accumulation can be used to eject a spin current into an adjacent layer (Edelstein effect³). Conversely, the injection of a spin current induces the associated spin polarization and charge current in the 2D states. Other types of conversion between charge and spin currents can also be obtained by SOC effects in three-dimensional (3D) conductors, namely the spin Hall effect of heavy metals⁴; however, the observed effects in two dimensions are considerably enhanced. Such spin–charge conversion phenomena have direct applications for spintronics technologies, which are based on the creation and detection of spin currents⁵. Given that the 3D spin Hall effect is already used in spintronics devices⁶, the observed effects in two dimensions offer much promise for device applications (Fig. 1, top).

The interplay between SOC and magnetism is of increasing importance. In conventional magnetic materials, ferromagnetic order, which results from exchange interaction, aligns neighbouring spins. A well-known consequence of SOC is magneto-crystalline anisotropy—the preferential alignment of electron moments along certain crystallographic directions ('easy axes'), via the coupling of electron motion to the crystalline lattice field. In systems that lack inversion symmetry, SOC induces

a chiral Dzyaloshinskii–Moriya interaction (DMI)^{7,8}, which takes the form:

$$H_{\text{DM}} = -(\mathbf{S}_1 \times \mathbf{S}_2) \cdot \mathbf{D}_{12} \quad (1)$$

Here \mathbf{S}_1 and \mathbf{S}_2 are neighbouring spins and \mathbf{D}_{12} is the Dzyaloshinskii–Moriya vector. The DMI is a chiral interaction that decreases or increases the energy of the spins depending on whether the rotation from \mathbf{S}_1 to \mathbf{S}_2 around \mathbf{D}_{12} is clockwise or anticlockwise. If \mathbf{S}_1 and \mathbf{S}_2 are initially parallel, then the effect of a sufficiently strong DMI (with respect to exchange and anisotropy) is to introduce a tilt around \mathbf{D}_{12} . DMI was initially understood as a super-exchange interaction in magnetic insulators^{7,8}, and later extended to non-centrosymmetric magnetic metals⁹. In a disordered magnetic alloy, a large SOC element could mediate such an interaction between two nearby magnetic atoms, with the resulting Dzyaloshinskii–Moriya vector being perpendicular to the plane formed by the three atoms. Crucially, this model was extended to magnetic multilayers, wherein inversion symmetry is broken by the presence of an interface¹⁰ (Fig. 1, bottom right). The existence of interfacial DMI was first demonstrated by the observation of spiral-like spatial modulations of the spin orientation with a winding periodicity related to the magnitude of the DMI (ref. 11). DMI also enables the formation of other chiral spin structures—in particular, chiral domain walls and skyrmions—that are possibly relevant to next-generation information storage devices (Fig. 1, bottom).

Recent developments in the techniques for thin-film growth and in the capabilities of *ab initio* calculations have enabled the synthesis of atomically flat surfaces and heterostructures, and the prediction of their electronic properties. A common thread across several such thin-film materials and heterostructures—heavy metal compounds and multilayers—is that the SOC strength at surfaces and interfaces is comparable to the other relevant energy scales, and so plays a pivotal part. In combination with surface and interface effects, this engenders fundamentally new spin-based phenomena that are robust to disorder and thermal fluctuations, with much promise for room-temperature spin-based applications.

Here we describe these diverse low-dimensional spin-based phenomena in the context of their SOC origin. We begin by detailing the progress on spin-polarized states at the surfaces of topological insulators, Rashba

¹Division of Physics and Applied Physics, School of Physical and Mathematical Sciences, Nanyang Technological University, 637371 Singapore. ²Data Storage Institute, A*STAR (Agency for Science, Technology and Research), 2 Fusionopolis Way #08-01, 138634 Singapore. ³Unité Mixte de Physique, CNRS, Thales, Université Paris-Sud, Université Paris-Saclay, Palaiseau 91767, France.

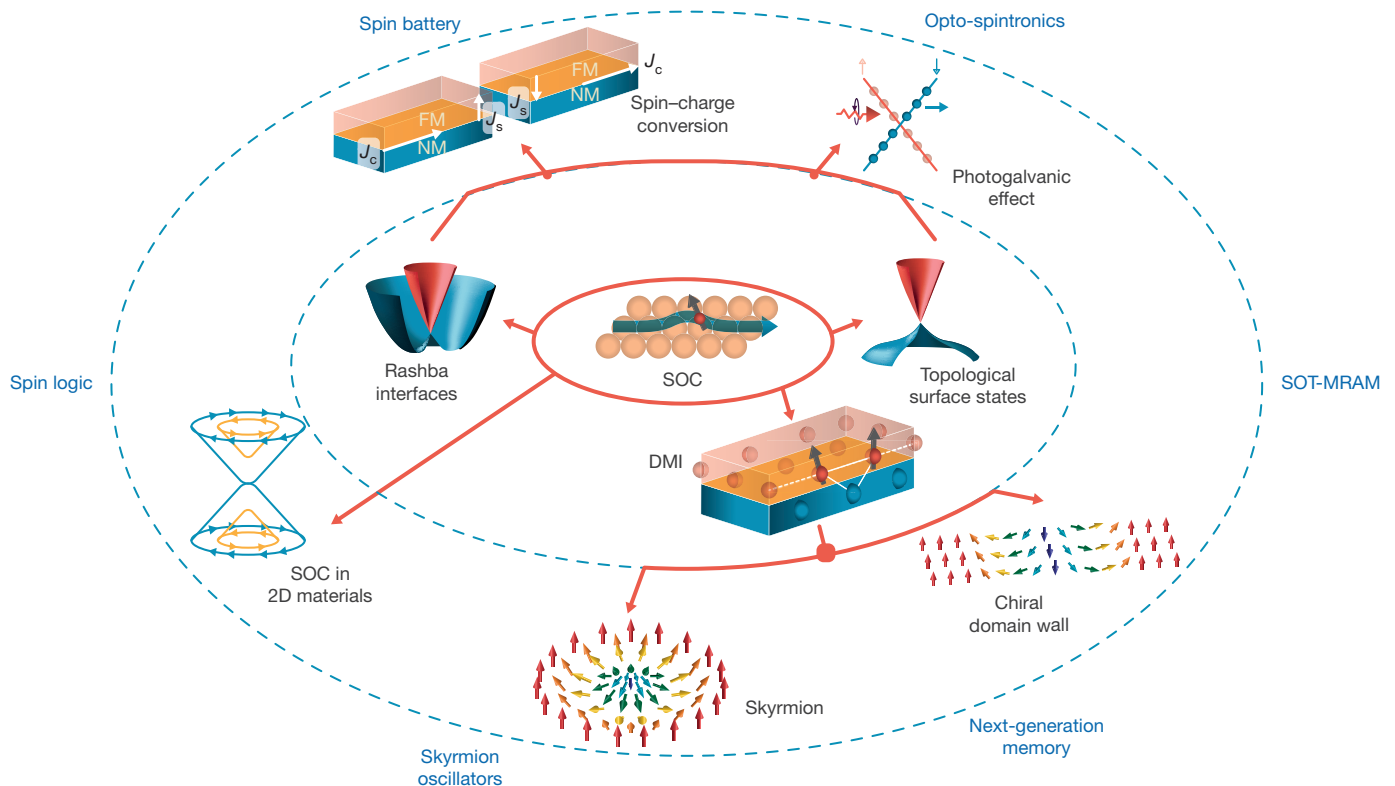


Figure 1 | Emergent phenomena from spin-orbit coupling (SOC) at surfaces and interfaces. A schematic illustration of the connection between the presence of strong SOC at material surfaces and interfaces (inner ellipse) and the resulting emergence of new interactions and electronic states (middle ellipse), such as Dzyaloshinskii–Moriya interaction (DMI; see Fig. 4a, e for details), Rashba interfaces (Fig. 2b, d) and topological surface

states (TSS; Fig. 2a, c). These emergent phenomena can in turn be used to generate new 2D spintronics effects (outer ellipse), such as spin–charge conversion (Fig. 2e, f and 3), the photogalvanic effect, enhanced SOC in 2D materials, such as graphene (Fig. 3d, e), magnetic skyrmions (Fig. 4b) and chiral domain walls (Fig. 4c), which have direct device applications (periphery). FM, ferromagnet; NM, non-magnetic material.

interfaces and atomically thin (2D) materials, and examine their utility towards the generation and conversion of spin currents. Next, we describe the developments on interfacial-DMI-induced non-collinear spin textures (skyrmions and chiral domain walls) in magnetic films, and techniques to generate, stabilize and manipulate them in devices. Finally, we explore the feasibility of realizing the technological promise of these diverse SOC-induced surface and interface phenomena towards room-temperature device applications.

Spin-polarized surface and interface states

Rashba states

The Rashba effect arises from SOC and broken inversion symmetry at material surfaces and interfaces¹, with the corresponding Hamiltonian:

$$H_R = v_0 \hat{z} \cdot (\mathbf{k} \times \boldsymbol{\sigma}) \quad (2)$$

Here v_0 is the Rashba parameter, $\boldsymbol{\sigma}$ is spin, \mathbf{k} is momentum and \hat{z} is the unit normal to the surface or interface. The Rashba effect results in spin-split 2D dispersion surfaces and, importantly, in the locking of spin and momentum degrees of freedom to each other (Fig. 2b).

Rashba SOC-split states have been investigated across various surfaces and interfaces^{2,12,13}, as shown for angle-resolved photoemission spectroscopy (ARPES) measurements of the Au(111) surface (Fig. 2d)¹³. Interface alloying of heavy elements with intermediate-weight metals can enhance the in-plane potential gradient via hybridization, leading to more pronounced Rashba effects, as on the Bi/Ag(111) alloyed interface ($v_0 = 3 \text{ eV \AA}$; ref. 14).

Topological surface states

In materials with heavy elements, strong SOC can split the p band by a large enough magnitude to flip the s – p band structure, inducing band

inversion. Notably, 2D heterostructures of $\text{Hg}_x\text{Cd}_{1-x}\text{Te}$ exhibit, in addition to such an inversion, an associated topological phase transition^{15,16}, which results in protected states at the edges of the sample. The presence of edge states in 2D heterostructures was subsequently generalized to 3D insulators^{17,18}. The surfaces or interfaces of such topological insulators must host protected states at time-reversal-invariant k -space points^{17,18}. These topological surface states have a nearly linear energy–momentum relationship (Fig. 2a)¹⁸. The Dirac Hamiltonian that describes these surface states, $H_D \propto v_0 \hat{z} \cdot (\mathbf{k} \times \boldsymbol{\sigma})$, has the same Rashba form (Equation (2)) and locks the spin and momentum degrees of freedom (Fig. 2a, c)¹⁸. However, whereas Rashba SOC leads to spin-split parabolic surface states in conventional metals, topological surface states are distinguished by their helical single Dirac cone character, which emerges from the requirement to connect the bulk valence and conduction bands.

ARPES measurements demonstrated the topological nature of surface states first in the indirect-bandgap semiconductor $\text{Bi}_{1-x}\text{Sb}_x$ (ref. 19) and then in a larger, direct-bandgap (300 meV) topological insulator Bi_2Se_3 (ref. 20). The discovery of a simple Dirac cone within the bandgap of bulk Bi_2Se_3 (Fig. 2c), with a chemical potential that is tunable via chemical doping²⁰ and the electric field effect²¹, has since led to the discovery of several other single-Dirac-cone topological insulators¹⁸.

The electronic transport of topological insulators is governed by the helical Dirac nature of topological surface states. First, surface-state transport arises from a 2D Dirac cone: therefore, it can be ambipolar, controlled by electric fields, and tuned through the Dirac point with a characteristic minimum conductivity²¹. Second, spin–momentum locking prevents backscattering between states of opposite momenta with opposite spins, as evidenced across several topological insulators²². Because backscattering dominates charge dissipation in conventional metals, quasiparticles of topological insulators are expected to exhibit

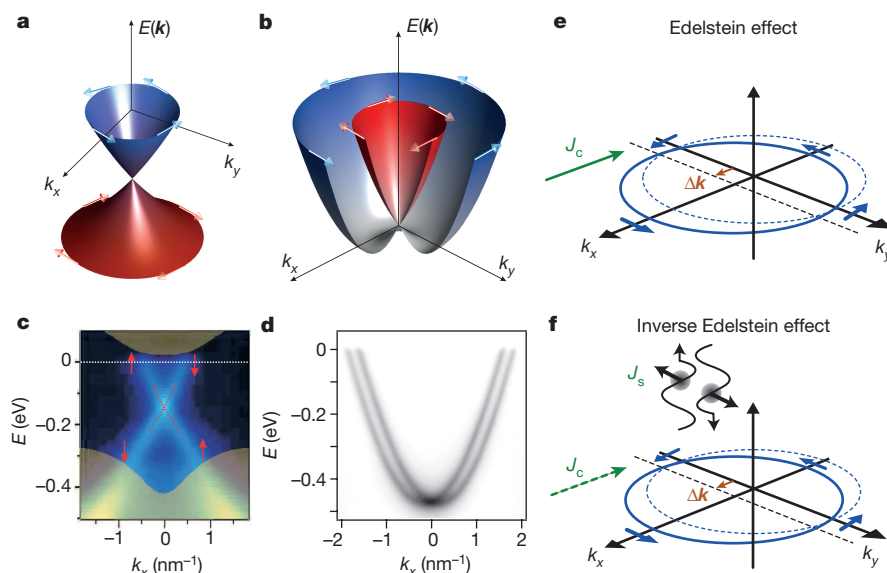


Figure 2 | Band structure and spin-charge conversion in spin-polarized 2D states. **a, b**, Schematic of the spin-polarized band structure (electron energy E as a function of in-plane momentum \mathbf{k}) of 2D electron states at the surfaces and interfaces of topological insulators (**a**) and at Rashba systems (**b**). The arrows indicate electron spin, with blue and red dispersion surfaces corresponding to opposite spin helicities. **c, d**, ARPES measurements of the 2D band structure, with respect to the Fermi level ($E = 0$), at the surface of the topological insulator $\text{Bi}_{2-x}\text{Ca}_x\text{Se}_3$ (**c**) and at the Rashba surface of $\text{Au}(111)$ (**d**). The red arrows in **c** indicate the spin orientation, and the bulk bands are schematized in brown. **e**, k -space schematic of charge-to-spin conversion in topological insulators via

the Edelstein effect. A charge current J_c at the surface causes a shift Δk of the Fermi contour, resulting in a non-zero spin density as a result of the helical spin orientation. This spin density can diffuse as a spin current in the adjacent material. **f**, Spin-to-charge conversion by the inverse Edelstein effect in topological insulators. Injecting a spin current (J_s ; spin-polarized wiggles) into the surface states of the topological insulator overpopulates states on one side of the Fermi contour and depopulates states on the other, generating a charge current. Panels **a** and **b** were adapted from figures 11.3(a) and 11.5(a), pages 181 and 183, of ref. 5, by permission of Oxford University Press. Panel **c** adapted from ref. 20, Nature Publishing Group. Panel **d** adapted from ref. 13, American Physical Society.

longer lifetimes, enabling ballistic phenomena such as the quantum Hall effect²³. Third, the polarization of light incident on a topological insulator can couple to the surface-state momentum, thereby generating spin-polarized photocurrents with high fidelity²⁴. Finally, magnetic doping of topological insulators breaks time-reversal symmetry, leading to a gap at the Dirac point²⁵ and a reorientation of the low-energy spin texture²⁶. The mediation of magnetism by Dirac fermions in topological insulators enables exotic phenomena, such as the quantum anomalous Hall effect²⁷.

Conversion between spin and charge currents

The helical spin polarization of Fermi contours of Rashba interfaces and surfaces of topological insulators enables the conversion between spin and charge currents by the Edelstein and inverse Edelstein effects³ (Fig. 2e, f). In a single-cone topological insulator, a charge current—that is, a shift of the Fermi contours in the direction of the electron motion (x) (Fig. 2e)—induces an overpopulation of spins in the transverse direction (y) as a result of spin-momentum locking, and is therefore associated with a non-zero spin accumulation. The spin accumulation can diffuse through an interface into an adjacent conducting material, resulting in a pure 3D spin current being injected into this material, without a net charge flow. The spin-charge conversion yield is quantified by the inverse length q_{ICS} —the ratio between the resultant 3D spin current density and the applied 2D charge current density²⁸. Furthermore, if the spin current is injected into a magnetic material, then the resulting spin-transfer torque can be used to switch its magnetization. Alternatively, an a.c. charge current could be used to induce ferromagnetic resonance.

In the inverse conversion of a spin current into a charge current by the inverse Edelstein effect on the interface of a topological insulator (Fig. 2f), the injection (extraction) of spins oriented along $+y$ ($-y$) into (from) the states of a helically spin-polarized Fermi contour populates (depopulates) states on the $+x$ ($-x$) side of the Fermi contour. This out-of-equilibrium distribution corresponds to a charge current. The conversion coefficient—the ratio between the induced 2D charge current density and applied 3D spin current density—is the inverse Edelstein effect length λ_{IEE}

(refs 29, 30). For a pure helical ground state $\lambda_{\text{IEE}} = v_F \tau$ (ref. 31), where v_F is the Fermi velocity and τ is the relaxation time of an out-of-equilibrium distribution in the interface states. The spin-charge conversion by the Edelstein and inverse Edelstein effects at Rashba interfaces can be described in an analogous fashion, but accounting for the partial compensation of the two Fermi contours of opposite chirality, yielding $\lambda_{\text{IEE}} = v_0 \tau / \hbar$ (refs 29, 30), where \hbar is the reduced Planck constant.

Examples of topological insulators in which charge-to-spin conversion effects have been observed include $(\text{Bi}_{1-x}\text{Sb}_x)_2\text{Te}_3$ (ref. 28) (Fig. 3a). Here, the charge-to-spin conversion of the applied a.c. charge current generates a vertical spin current, and a spin-transfer torque on the magnetization of the deposited NiFe layer. When tuning the Fermi level across the Dirac point by varying the doping x , the conversion coefficient q_{ICS} does not change sign because both charge and spin chiralities change simultaneously (Fig. 3a). However, q_{ICS} exhibits a sharp minimum at the Dirac point, at which a finite scattering rate between the upper and lower cone can mix the spin polarizations. Similar results of the generation of spin polarization and spin torques have been obtained for other topological insulators^{32–35}.

Spin-to-charge conversion on topological insulators by the inverse Edelstein effect has been achieved both in spin pumping³⁶ and in tunneling spin injection³⁵ experiments. In the case of spin pumping conversion (Fig. 3b, c)³¹, a vertical spin current J_s , which is produced by ferromagnetic resonance from a ferromagnetic layer (NiFe), is injected into the topological insulator α -Sn through an intermediate Ag layer (Fig. 3c). The persistence of the Dirac cone on α -Sn after Ag deposition is verified by ARPES (Fig. 3b), whereas direct deposition of Fe on Sn destroys the Dirac cone. A charge current induced by the inverse Edelstein effect is consistently found with Ag on α -Sn (Fig. 3c, blue), and not with Fe (Fig. 3c, green). Here, $\lambda_{\text{IEE}} = 2.1$ nm at room temperature—a spin-charge conversion efficiency that is at least an order of magnitude higher than that obtained with the inverse spin Hall effect in metals such as Pt or W (ref. 31).

Analogous spin-charge conversions by the Edelstein and inverse Edelstein effects can also be obtained at Rashba interfaces^{29,37}. However,

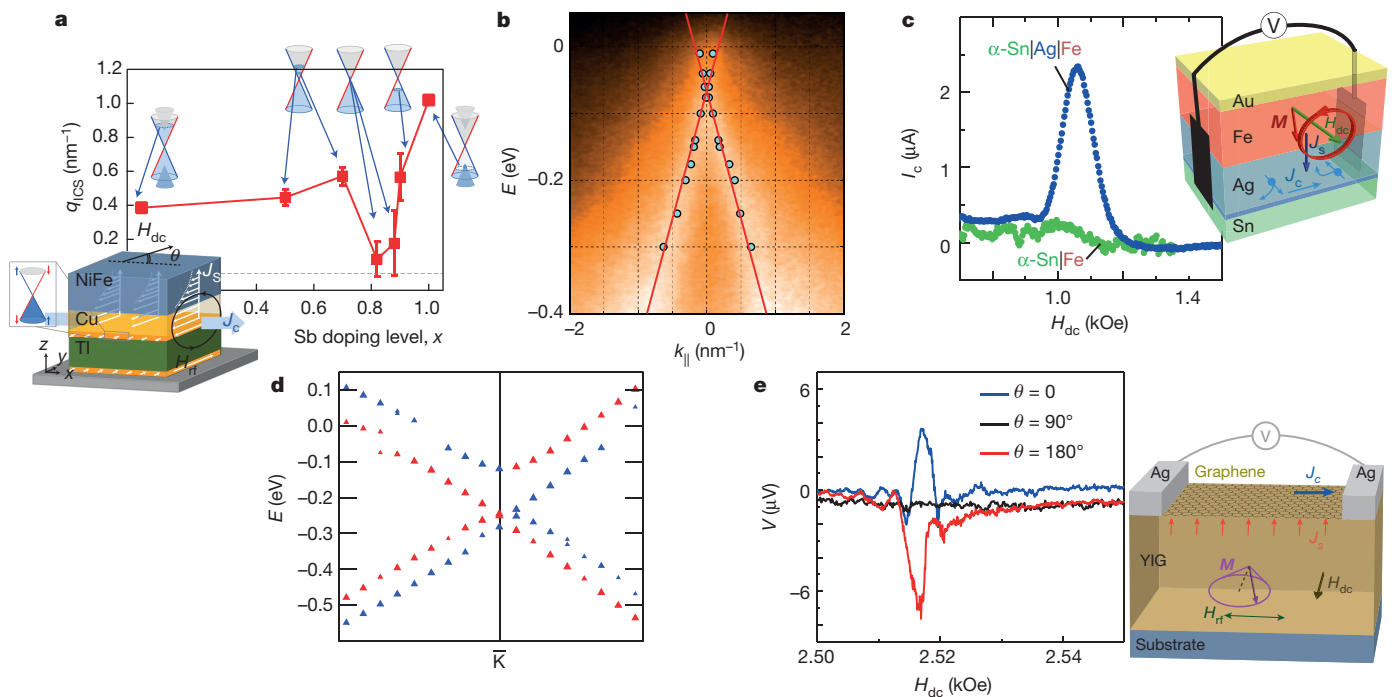


Figure 3 | Spin-charge conversion experiments. **a**, Charge-to-spin conversion by the Edelstein effect in spin torque ferromagnetic resonance (ST-FMR) experiments on the topological insulator $(\text{Bi}_{1-x}\text{Sb}_x)_2\text{Te}_3$ at $T = 10$ K. The graph shows the spin-charge conversion coefficient q_{ICS} for several Sb concentrations x , and the corresponding Fermi levels in the Dirac cone. (Error bars represent the standard deviation over five samples with different dimensions.) Inset, schematic of the device, in which an applied longitudinal a.c. charge current I_c (associated with an a.c. magnetic field H_{rf}) is converted by the Edelstein effect into a vertical spin current J_s . This spin current is injected through a Cu layer into the top NiFe layer and detected by a ST-FMR-induced d.c. voltage, in an external d.c. field H_{dc} applied at an angle θ . **b**, **c**, Spin-to-charge conversion by the inverse Edelstein effect in spin-pumping experiments on the topological insulator $\alpha\text{-Sn}$. **b**, ARPES intensity maps (for varying electron energy E with respect to the Fermi level and in-plane momentum k_{\parallel}) show that the surface-state Dirac cone on $\alpha\text{-Sn}$ is preserved (symbols on the red lines represent maxima in ARPES intensity scans) even after Ag deposition (to a thickness

of 2.3 nm). **c**, A d.c. voltage generated by the inverse Edelstein effect is observed when $\alpha\text{-Sn}$ is covered by Ag (blue), but not when it is directly covered by Fe, which destroys the Dirac cone (green). Inset, a FMR spin-pumping device (the magnetization M of the ferromagnetic layer in an external field $B = \mu_0 H_{\text{dc}}$, where μ_0 is the permeability of free space, excited by an a.c. field) injects a vertical spin current J_s through an Ag layer into the surface states of the topological insulator, which generates a charge current I_c . **d**, Calculated band structure of graphene after Au intercalation between graphene and substrate, matched to ARPES measurements near the \bar{K} point. Blue and red symbols indicate opposite spin orientations derived from fits to spin-resolved ARPES measurements. **e**, Spin-to-charge conversion by spin pumping from yttrium iron garnet (YIG) into graphene: experimental set-up (inset) and lateral voltage V induced by the inverse Edelstein effect for opposite applied fields. Panel **a** adapted from ref. 28, Nature Publishing Group. Panels **b** and **c** adapted from ref. 31, American Physical Society. Panel **d** adapted from ref. 47, Nature Publishing Group. Panel **e** adapted from ref. 50, American Physical Society.

owing to the compensation between the two Fermi contours of the Rashba interface, the conversion is generally not as efficient as for topological insulators. An exception is the interface between the insulating oxides SrTiO_3 and LaAlO_3 , for which inverse Edelstein effect values as large as 6 nm have been reported³⁸.

Materials directions and applications

Despite tremendous excitement surrounding topological insulators, the development of technologically relevant materials has been hindered by several issues. In binary topological insulators (for example, Bi_2Se_3), intrinsic defects pin the chemical potential within the bulk bands, diminishing the contribution of the surface-state transport²⁰. Although chemical doping could tune the chemical potential into the bandgap, this markedly reduces the mean free path²². Therefore, recent efforts have used epitaxial techniques to fabricate ternary topological insulators with better control over the chemical potential and mobility³⁹. Other efforts towards functional topological insulators include systems with larger Rashba SOC and stack engineering of strong SOC materials⁴⁰. Another emerging direction lies at the crossover between Dirac and Rashba behaviour, with the Dirac/Rashba character of quasiparticles tuned by film thickness (for example, few-layer Bi_2Se_3 films⁴¹) or chemical potential (for example, Sb; ref. 42).

The partial freezing of backscattering for surface and interface states of topological insulators, which results in reduced energy dissipation

by electrical currents, is a notable advantage for their use in low-power nano-devices. As a next step, the ambipolarity of topological surface states could be exploited to create, for example, topological p-n junctions⁴³ and spin transistors. The spin Hall effect of heavy metals is already used in spin-orbit torque magnetic random-access memory (SOT-MRAM) switching⁶, and similar conversions by topological insulators or Rashba interfaces are expected to be much more energy-efficient³².

Enhanced SOC in 2D materials

As with surface and interface states of topological insulators, the 2D electronic states of single-layer graphene are characterized by linearly dispersing Dirac cones at the K and K' points in reciprocal space⁴⁴. In contrast to topological insulators, the SOC magnitude for the sp^2 bonded structure of pristine graphene is small (about $10\text{ }\mu\text{eV}$)⁴⁴, and Dirac cones in graphene are generally supposed to be spin-degenerate. However, SOC can be greatly enhanced by proximity and hybridization with adjacent materials. In-plane and out-of-plane deformations that mix the sp^2 and sp^3 orbitals in strained or buckled graphene, as well as interactions with adatoms and electric fields, can also lead to enhancements in SOC^{45–49}.

The SOC effects of an ordered interface between graphene and another material on the energy dispersion of graphene can be expressed as⁴⁸

$$E_{\alpha\beta} = \frac{\alpha\lambda}{2} + \beta \left(k^2 + \frac{\lambda^2}{4} \right)^2$$

with $\beta = +1$ (-1) for the upper (lower) dispersion surfaces, $\alpha = +1$ (-1) for the split upper (lower) dispersion surfaces, and where λ is the SOC constant of the system and $k = |\mathbf{k}|$. Therefore, SOC-induced splitting of graphene bands leads further to spin–momentum locking and chiral spin orientations; for example, the large SOC splitting (about 100 meV) for graphene on Au (see, for example, Fig. 3d) is caused by the strong hybridization between the Dirac-cone states and d states of Au (ref. 47).

Enhancement of SOC in graphene can similarly lead to 2D spin–charge conversion effects. For example, ferromagnetic-resonance spin pumping from yttrium iron garnet (YIG) into graphene induces a definite broadening and a lateral voltage characteristic of the inverse Edelstein effect (Fig. 3e)⁵⁰ that is ascribed to the SOC induced by proximity with YIG. Here, $\lambda_{\text{IEE}} = 10^{-3}$ nm, which is much smaller than with topological insulators or Rashba interfaces, indicating a moderate SOC splitting. On the other hand, intercalation of, for example, an Au layer may further enhance SOC (Fig. 3d), leading to more efficient conversion. Similarly, non-local voltages, corresponding to enhanced SOC, have been observed on graphene decorated with small amounts of adatoms or nanoparticles⁵¹, and attributed to skew scattering on enhanced SOC perturbations⁴⁹.

Furthermore, it is appealing to explore other 2D systems that possess larger intrinsic SOC, such as layered transition-metal dichalcogenides (TMDs)⁵². Spin pumping and spin-transfer torque experiments on TMDs^{53,54}, notably MoS₂ and WS₂, have shown promise for spin–charge conversion. Whereas semiconducting TMDs have limited utility for spintronics, owing to their small mobility and large resistivity, heterostructures of large-SOC TMDs and high-mobility graphene might be more useful. Considering the large SOC enhancement of graphene in proximity with WS₂ (ref. 54), larger spin–charge conversion effects can be expected in such structures. Heterostructures of the topological insulator Bi₂Te₂Se and graphene have also shown efficient transfer of spin current to the graphene layer⁵⁵.

Interfacial spin interactions and chiral magnetism

Interfacial DMI

Magnetic materials that lack inversion symmetry can host the DMI in the presence of strong SOC. Consequently, neighbouring spins tilt with respect to each other, leading to spatial modulations of the spin orientation. If the magnitude of \mathbf{D}_{12} (Equation (1)) is sufficiently large, then the competition between the ‘winding’ DMI and ‘aligning’ exchange interactions can give rise to non-collinear ground states^{11,56–58}. Such chiral spin structures were initially identified in non-centrosymmetric single crystals^{56,58–60}. However, of increasing scientific interest and technological relevance, is their manifestation in films and multilayers with interfacial DMI (Fig. 4a).

Interfaces between ultrathin magnetic materials and metals with strong SOC can host DMI, owing to broken inversion symmetry¹⁰. Large interfacial DMI (effective magnitude of \mathbf{D}_{12} comparable to the exchange constant J) and ensuing spin textures were first observed in ultrathin epitaxial magnetic films: spin spirals in Mn on W(110) (ref. 11), and skyrmions on Fe and Fe/Pd on Ir(111) (refs 61, 62). In the presence of out-of-plane anisotropy (K), this also leads to collinear magnetic domains separated by Neel-type chiral domain walls, that is, those with the chirality of spin rotation through the domain wall determined by the orientation of \mathbf{D}_{12} (Fig. 4c)^{63–66}. More recently, it has been possible to demonstrate the presence of DMI-induced skyrmions at room temperature in magnetic multilayers grown by sputtering^{67–69}, opening the possibility for using such structures in spintronics technologies⁷⁰.

The schematic in Fig. 4a (inset) shows such an interface between a magnetic film (Co) and a metal with strong SOC (Pt). *Ab initio* studies have provided insights into the mechanism for interfacial DMI across several materials^{61,71,72}. For instance, it has been shown that, at a

Co/Pt interface, the DMI is strongest in the Co layer that is closest to the interface, and relatively negligible in other Co layers (Fig. 4a, bottom)⁷². Furthermore, the energy source of the large DMI in the interface Co layer is located predominantly in the neighbouring large-SOC Pt layer (Fig. 4a, top), adding credence to a direct correspondence between the DMI and interfacial electronic states. The DMI in such multilayers has been measured directly using spin-wave dispersion mapping techniques, and recent studies have shown its inverse relationship with the thickness of the magnetic layer—a clear signature of its interfacial nature⁷³.

Magnetic skyrmions in thin films

The large DMI observed in these magnetic multilayers is of particular interest because it induces new chiral spin textures, known as magnetic skyrmions. Skyrmions are quasi-2D spin textures wherein the out-of-plane magnetization is inverted at the centre and rotates smoothly with a fixed chirality across its width (Fig. 4b). Skyrmions are distinguished by the topological number S that characterizes the winding of the normalized local magnetization $\mathbf{m}(\mathbf{r})$ (at position \mathbf{r}) in 2D systems:

$$S = \frac{1}{4\pi} \int \mathbf{m}(\mathbf{r}) \cdot (\partial_x \mathbf{m}(\mathbf{r}) \times \partial_y \mathbf{m}(\mathbf{r})) d\mathbf{r} = \pm 1$$

This topological number indicates that the skyrmion magnetization covers the entirety (4π) of the unit sphere⁶⁰. Even though skyrmion-like objects (such as magnetic bubbles) can be stabilized without DMI (for example, by dipolar interactions), they would not have a fixed topological number S , which is crucial to skyrmion properties^{57,60}. In fact, the non-trivial topology of the skyrmion affords an energy barrier, protecting its spin structure: the spin configuration cannot be twisted continuously to obtain a different S (for example, that of uniform polarization). Another key property of the magnetic skyrmion is its solitonic nature with a finite extension: it can move as a particle for as long as it can be stabilized in a uniform ferromagnetic background.

Skyrmions in ultrathin magnetic films were first observed in epitaxial magnetic monolayers on heavy metal substrates (Fe on Ir(111)), for which D_{12}/J can be extremely large (about 1)^{61,62}. Here, skyrmions form a stable lattice configuration and the large value of D_{12}/J results in spin rotation over shorter length scales, reducing the skyrmion size to just a few atomic lengths. However, the formation of skyrmions herein requires low temperatures (about 10 K). Isolated skyrmions can be stabilized in a metastable state by applying a magnetic field, for example in Fe/Pd on Ir(111) (Fig. 4d)⁶², or at zero field if the DMI is sufficiently large, yet smaller than the threshold value for stabilizing a skyrmion lattice or a spin spiral^{74,75}, that is, if $D < 4\sqrt{AK}/\pi$, where D is the normalized DMI per unit area of the film, A is the exchange stiffness and K characterizes the out-of-plane anisotropy. Such individual skyrmions can further be nucleated or annihilated (Fig. 5a). Owing to their topological protection, they are highly stable⁷⁶, but only in applied field and at low temperature, which limits their use as individual particles in devices.

A crucial challenge for device applications is the room-temperature stabilization of such small individual skyrmions. The Néel or Curie temperatures of bulk materials that are known to host skyrmions are generally below room temperature. In addition, skyrmions in ultrathin magnetic monolayers (see, for example, Fig. 4d) are stable only at low temperatures. A prominent direction towards practical room-temperature skyrmions is the development of multilayers with additive interactions at successive interfaces^{67,68}. First, the interfacial DMI can be enhanced by an appropriate choice of elements forming the multilayer stack. For example, in the case of Ir/Co/Pt multilayers, the sign of D is opposite for Ir/Co and Pt/Co interfaces. As a result, when Ir and Pt layers are on opposite sides of the Co layer, their effects are additive, thus increasing the net DMI magnitude⁶⁷. Furthermore, such trilayer stacks can be repeated (for example, the ten repeats of (Ir/Co/Pt); Fig. 4e), and skyrmions in successive stacks are coupled through ultrathin non-magnetic layers, leading to columnar room-temperature skyrmions (≥ 30 nm) stabilized by their larger magnetic volume^{67,68}.

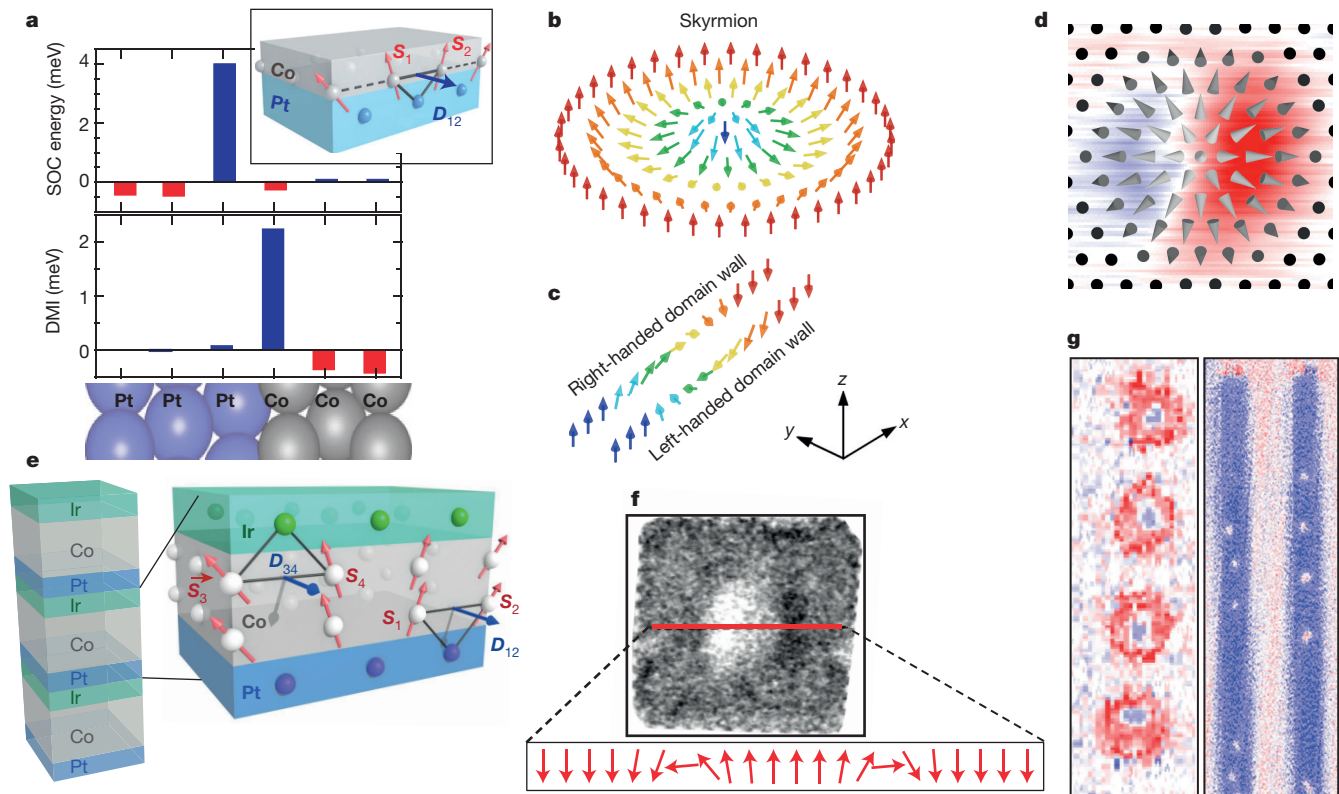


Figure 4 | Interfacial DMI and chiral spin textures. **a**, Anatomy of interfacial DMI from *ab initio* calculations. Bottom, Layer-resolved DMI in a Pt/Co bilayer. Top, distribution of SOC energies associated with the DMI in the interfacial Co layer. Inset, a schematic of DMI at the interface between a ferromagnetic metal with out-of-plane magnetization (Co, grey) and a strong SOC metal (Pt, blue). The DMI vector D_{12} , associated with the triangle composed of two Co atoms and a Pt atom, is perpendicular to the plane of the triangle. $S_{1,2}$, neighbouring spins. **b**, **c**, Schematics of the spin configuration in interfacial-DMI-induced chiral spin textures such as magnetic skyrmions (**b**) and chiral Néel domain walls (**c**), with the colour scale corresponding to the out-of-plane magnetization component. **d**, SP-STM imaging of an individual skyrmion (with a diameter of 8 nm at a field of 3.25 T) in a Fe/Pd bilayer on Ir(111), acquired in constant-current topographic mode, with an in-plane magnetized tip, with the modelled magnetization overlaid (arrows). **e**, Skyrmion stabilization in multilayers,

illustrated using a multilayer stack of Ir/Co/Pt. The close-up of the trilayer shows DMI vectors (D_{12} and D_{34}) at the top (Co/Ir) and bottom (Pt/Co) interfaces of Co. The effective DMI magnitude is enhanced by the same direction of D_{12} and D_{34} at the different interfaces. **f**, Room-temperature skyrmions in a Pt/Co/MgO multilayer in a lithographed 400 nm × 400 nm square, seen by XMCD-PEEM, with the magnetization profile along the red line shown below. **g**, Room-temperature skyrmions in (Ir/Co/Pt) × 10 multilayers patterned into 300-nm-diameter disks (left) or 200-nm-wide tracks (right), seen by STXM. Panel **a** (main panel) adapted from ref. 72, American Physical Society. Panel **a** (inset) adapted from ref. 70, Nature Publishing Group. Panel **d** reproduced from ref. 62, American Association for the Advancement of Science. Panels **e** and **g** adapted from ref. 67, Nature Publishing Group. Panel **f** adapted from ref. 78, Nature Publishing Group.

Room-temperature skyrmions have also been found in magnetic bilayers (Figs. 4f, 5b), although generally with a larger diameter^{69,77–79}. These efforts offer promising directions towards stack engineering of magnetic interactions to tune skyrmion properties in films for device applications⁸⁰.

Detection and manipulation of chiral spin textures

Skyrmions in epitaxial films were first imaged using spin-polarized scanning tunnelling microscopy (SP-STM; Fig. 4d)^{61,62}. Since then, they have been imaged in sputtered multilayer films using various magnetic microscopy techniques, including scanning transmission X-ray microscopy (STXM; Fig. 4g)^{67,68}, photoemission electron microscopy (PEEM; Fig. 4f)⁷⁸, spin-polarized low-energy electron microscopy (SPLEEM)⁷⁷, and magneto-optical Kerr effect (MOKE) microscopy (Fig. 5b)⁶⁹. Importantly, skyrmions can also be detected using a variety of thermodynamic and transport techniques⁸¹. In particular, the Berry phase that is accumulated by electrons traversing the 2D spin texture of skyrmions results in an additional component in anomalous Hall effect measurements, known as the topological Hall effect^{60,81}. The Hall signal can be used to detect the presence of skyrmions and to address their motion in films and devices^{81,82}. However, such Hall signatures of skyrmions have been detected thus far only in bulk crystal and films with intrinsic

DMI^{81–83}; these techniques remain to be established in multilayer films with interfacial DMI.

Magnetic skyrmions, owing to their small size and non-trivial topology, are attractive candidates for data storage in magnetic materials—provided that they can be nucleated, moved and read. Several nucleation techniques have been explored with micromagnetics simulations^{75,84}. In SP-STM experiments on Fe/Pd bilayers (Fig. 5a), individual skyrmions were nucleated and deleted using the current injected from the STM tip⁶². In other experiments, skyrmions have been created by applying field pulses⁶⁸. A remarkable result in this regard is the recent demonstration of “blowing of skyrmion bubbles”^{69,85}, generated by the current divergence out of a constriction (Fig. 5b). In future, skyrmions should be able to be moved with notable ease compared with, for example, domain walls⁸² by exploiting the SOT provided by the spin current^{75,86,87}, which emerges naturally from the spin Hall effect of the neighbouring heavy metal layers. The dynamic properties of skyrmions have been explored using micromagnetics simulations and microscopy techniques in device configurations^{68,69}. These works demonstrate that skyrmions can be manipulated with current and field pulses in lithographed geometric structures (Fig. 5b, c)^{68,69}—techniques that can be incorporated in memory devices with relative facility.

These properties of magnetic skyrmions portend great potential towards realizing high-density and energy-efficient memory^{86,87}. Several

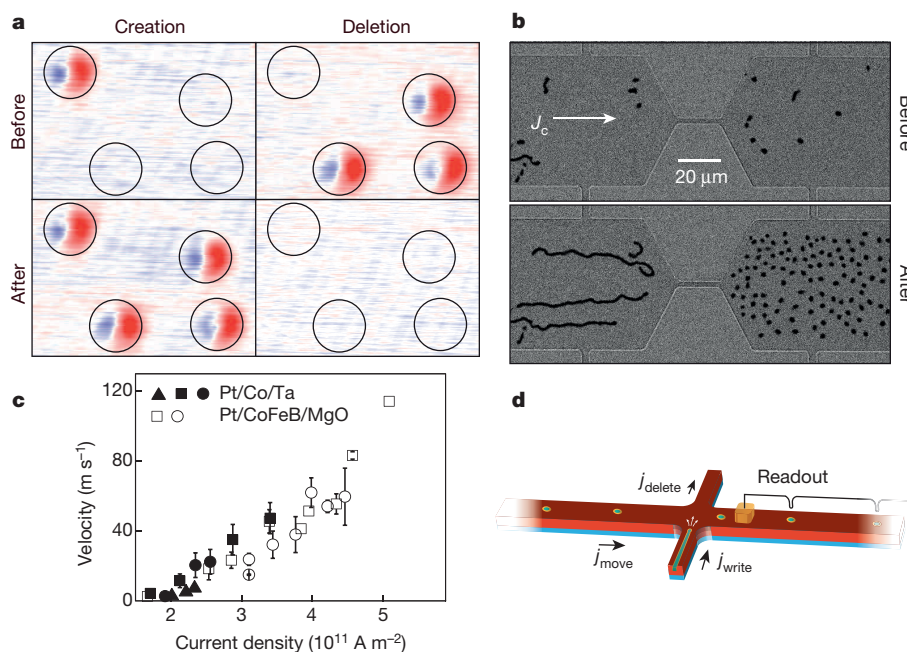


Figure 5 | Manipulation of magnetic skyrmions. **a**, Individual skyrmions (with diameters of 8 nm at a field of 3.25 T) in Fe/Pd/Ir(111) before and after SP-STM manipulation, demonstrating the creation and annihilation of individual skyrmions at specific locations. **b**, Skyrmions in a Ta/CoFeB/TaO_x structure, before (top) and after (bottom) applying a current pulse through a constriction, with current-induced nucleation and subsequent motion of several skyrmions, as seen by magneto-optical Kerr effect (MOKE) microscopy. **c**, Experimental measurement of the current-induced skyrmion velocity in tracks of Pt/Co/Ta and Pt/

CoFeB/MgO (different symbols represent results from different devices) multilayers using STXM. (Error bars denote the standard deviation of multiple measurements.) **d**, Schematic of a skyrmion-based memory device in which skyrmions could be deleted, moved and written by the corresponding current j . Panel **a** adapted from ref. 62, American Association for the Advancement of Science. Panel **b** adapted from ref. 69, American Association for the Advancement of Science. Panel **c** reproduced from ref. 68, Nature Publishing Group. Panel **d** adapted from ref. 100, American Association for the Advancement of Science.

applications and architectures have been proposed and modelled, including skyrmion-based memory devices (Fig. 5d) analogous to domain-wall-based racetrack memory⁸⁸. The interest in skyrmions with respect to domain walls is the smaller current that is needed for their displacement and the weaker influence of defects on skyrmion motion⁷⁵. More efficient SOT could also be obtained by using the spin–charge conversion at the interfaces of topological insulators instead of the spin Hall effect.

DMI of intermediate strength has direct relevance to chiral domain walls⁶³, which are also being actively explored^{165,66,89}. The motion of magnetic domain walls under SOT depends on the relative configuration of the domain-wall magnetization and the type of SOT under consideration⁹⁰. The large domain-wall velocity observed in perpendicular anisotropy films deposited on heavy metals was initially understood as emerging from the SOT at the interface between the magnet and the heavy metal⁹¹. However, it has been recently understood that the marked enhancement in domain-wall velocity results from the stabilization of chiral, Néel-type domain walls by interfacial DMI⁹². This DMI-induced stabilization suppresses the Walker breakdown mechanism that typically limits domain-wall dynamics and explains the efficient SOT action on this type of domain wall⁶⁶. In addition, the domain-wall chirality, which is determined by DMI, corresponds to a fixed direction of motion; hence, all chiral domain walls move in the same direction in a given stack structure.

The motion of DMI-stabilized chiral domain walls can lead to new realizations of nanoscale data storage. For example, consider a memory element that stores the information by using the domain-wall position. The domain wall can have two stable positions (for example, using a notch along a short stripe). Here, the magnetization switch can be measured using a magnetic tunnel junction⁹³. Such memory architectures based on switching nanoscale spin structures require much less current than does conventional MRAM, wherein the magnetization of the entire device needs to be reversed. Furthermore, the fast SOT-induced motion of chiral domain walls⁸⁹ is relevant to advancing the development of other concepts, such as domain-wall racetrack memory⁸⁸.

Applications and outlook

The interplay between SOC and inversion-symmetry breaking has given rise to fascinating phenomena at surfaces and interfaces, especially in the past decade. Topological insulators have been described as a “new state of quantum matter”¹⁹, and the emergence of interfacial DMI has given rise to non-trivial spin structures. However, the remarkable properties induced by SOC at surfaces and interfaces go beyond chiral magnets and spin topology, including several types of 2D materials with intrinsic or engineered SOC. The emergent characteristics of these SOC-induced phenomena, which are robust at room temperature, offer several potential applications.

First, spin–momentum locking in topological insulators can be exploited, via their interaction with normal metals, to obtain unprecedented efficiency in spin–charge conversions. High spin–charge conversion efficiency will probably be harnessed in future spintronic devices, such as SOT-MRAM or nano-batteries. Second, the protection from backscattering in topological insulators can be used in low-dissipation devices. The topologically protected skyrmion spin configurations ultimately represent the smallest achievable size for an emergent non-volatile magnetic memory element in magnetic films, with immediate relevance to information storage. Skyrmions can be moved, created and annihilated in nanostructures, making them suitable for ‘abacus’-type applications such as racetrack memory.

Although several of these avenues for application have emerged only in the past decade, the rapid advances along this front make us optimistic about the time frame in which we can reasonably expect to see devices that realize the potential of SOC-induced properties. Consider skyrmions in magnetic multilayers: pioneering efforts on epitaxial films at low temperature^{61,62} were soon followed by room-temperature observations in the kind of sputtered multilayer films^{67–69,78} typically used in spintronics technologies. Demonstrations of their small size, electrical nucleation and motion—all under ambient conditions—offer further technological promise. Other devices include skyrmion-based transistors⁹⁴, oscillators and microwave detectors⁹⁵. Furthermore, concepts of SOT-MRAM could be extended to utilize skyrmions.

The applications that could emerge from topological insulators may evolve over a range of time frames. Fundamental materials challenges remain to be overcome, including the integration of topological-insulator compounds into existing elements of spintronic technologies and the control of defects that impede their exotic properties at room temperature. However, recently discovered topological insulators demonstrate unprecedented spin–charge conversion efficiency at room temperature³¹. They are therefore suitable candidates to replace spin-Hall-effect heavy metals—for example, for memory state switching of SOT-RAM—in the near future³². Another exciting prospect is the use of topological insulators to generate efficient SOT for manipulating skyrmions and chiral domain walls. Charge-to-spin conversion by Bi/Ag Rashba interfaces has been proposed as a way to develop non-volatile and complementary metal–oxide–semiconductor (CMOS)-compatible spin logic devices⁹⁶, and spin-to-charge conversion in ferromagnet–topological-insulator devices can be used for microwave-driven spin batteries⁹⁷ and electrical power generators that exploit spin currents induced by temperature gradients⁹⁸. Furthermore, the high fidelity of room-temperature spin-polarized photocurrents generated on topological insulators²⁴ offer promising opto-spintronics applications, including transparent conducting electrodes, phase modulators and polarizers. Finally, upon overcoming the materials challenges, we foresee that the ambipolarity of topological surface states could be used to create low-power spintronics devices such as a topological p–n junction⁴³. One may envisage a device where an applied magnetic field turns such a p–n junction into an electronic Mach–Zender interferometer⁹⁹, enabling a tunable junction transmission with spin filtering properties.

The discovery of the novel states induced by SOC and inversion-symmetry breaking at surfaces and interfaces opens up such a broad perspective that the introduction of their topological properties will have a definitive and substantial effect on the technology of spintronics.

Received 16 October 2015; accepted 2 August 2016.

1. Rashba, E. I. Properties of semiconductors with an extremum loop. I. Cyclotron and combinational resonance in a magnetic field perpendicular to the plane of the loop. *Sov. Phys. Solid State* **2**, 1109–1122 (1960).
2. Manchon, A., Koo, H. C., Nitta, J., Frolov, S. M. & Duine, R. A. New perspectives for Rashba spin–orbit coupling. *Nat. Mater.* **14**, 871–882 (2015).
3. Edelstein, V. M. Spin polarization of conduction electrons induced by electric current in two-dimensional asymmetric electron systems. *Solid State Commun.* **73**, 233–235 (1990).
This paper describes the mechanism that leads to spin polarization when a charge current passes through a 2D electronic system with SOC, which is the basis of the spin–charge conversion mechanism described here.
4. Hoffmann, A. Spin Hall effects in metals. *IEEE Trans. Magn.* **49**, 5172–5193 (2013).
5. Valenzuela, S. O. in *Spin Current* (eds Maekawa, S. et al.) Ch. 11 (Oxford Univ. Press, 2012).
6. Cubukcu, M. et al. Spin–orbit torque magnetization switching of a three-terminal perpendicular magnetic tunnel junction. *Appl. Phys. Lett.* **104**, 042406 (2014).
7. Dzyaloshinsky, I. A thermodynamic theory of ‘weak’ ferromagnetism of antiferromagnetics. *J. Phys. Chem. Solids* **4**, 241–255 (1958).
8. Moriya, T. Anisotropic superexchange interaction and weak ferromagnetism. *Phys. Rev.* **120**, 91–98 (1960).
9. Fert, A. & Levy, P. M. Role of anisotropic exchange interactions in determining the properties of spin-glasses. *Phys. Rev. Lett.* **44**, 1538–1541 (1980).
10. Fert, A. Magnetic and transport properties of metallic multilayers. *Mater. Sci. Forum* **59–60**, 439–480 (1990).
This work provides a theoretical description of the mechanism that underlies DMI generated at the interface between a ferromagnetic layer and a strong SOC metal.
11. Bode, M. et al. Chiral magnetic order at surfaces driven by inversion asymmetry. *Nature* **447**, 190–193 (2007).
This paper represents the first identified observation of interfacial DMI-induced spin modulations in ultrathin magnetic films.
12. LaShell, S., McDougall, B. A. & Jensen, E. Spin splitting of an Au(111) surface state band observed with angle resolved photoelectron spectroscopy. *Phys. Rev. Lett.* **77**, 3419–3422 (1996).
13. Nechaev, I. A. et al. Hole dynamics in a two-dimensional spin–orbit coupled electronic system: theoretical and experimental study of the Au(111) surface state. *Phys. Rev. B* **80**, 113402 (2009).
14. Ast, C. R. et al. Giant spin splitting through surface alloying. *Phys. Rev. Lett.* **98**, 186807 (2007).
15. Bernevig, B. A., Hughes, T. L. & Zhang, S.-C. Quantum spin Hall effect and topological phase transition in HgTe quantum wells. *Science* **314**, 1757–1761 (2006).
16. König, M. et al. Quantum spin Hall insulator state in HgTe quantum wells. *Science* **318**, 766–770 (2007).
17. Fu, L., Kane, C. L. & Mele, E. J. Topological insulators in three dimensions. *Phys. Rev. Lett.* **98**, 106803 (2007).
18. Hasan, M. Z. & Kane, C. L. *Colloquium: Topological insulators.* *Rev. Mod. Phys.* **82**, 3045–3067 (2010).
19. Hsieh, D. et al. A topological Dirac insulator in a quantum spin Hall phase. *Nature* **452**, 970–974 (2008).
20. Hsieh, D. et al. A tunable topological insulator in the spin helical Dirac transport regime. *Nature* **460**, 1101–1105 (2009).
A single-Dirac-cone topological insulator with spin-momentum locking at room temperature and a tunable chemical potential is demonstrated for the first time.
21. Checkelsky, J. G., Hor, Y., Cava, R. J. & Ong, N. Bulk band gap and surface state conduction observed in voltage-tuned crystals of the topological insulator Bi₂Se₃. *Phys. Rev. Lett.* **106**, 196801 (2011).
22. Beidenkopf, H. et al. Spatial fluctuations of helical Dirac fermions on the surface of topological insulators. *Nat. Phys.* **7**, 939–943 (2011).
23. Brüne, C. et al. Quantum Hall effect from the topological surface states of strained bulk HgTe. *Phys. Rev. Lett.* **106**, 126803 (2011).
24. McIver, J. W., Hsieh, D., Steinberg, H., Jarillo-Herrero, P. & Gedik, N. Control over topological insulator photocurrents with light polarization. *Nat. Nanotechnol.* **7**, 96–100 (2011).
25. Chen, Y. L. et al. Massive Dirac fermion on the surface of a magnetically doped topological insulator. *Science* **329**, 659–662 (2010).
26. Xu, S.-Y. et al. Hedgehog spin texture and Berry’s phase tuning in a magnetic topological insulator. *Nat. Phys.* **8**, 616–622 (2012).
27. Chang, C.-Z. et al. Experimental observation of the quantum anomalous Hall effect in a magnetic topological insulator. *Science* **340**, 167–170 (2013).
28. Kondou, K. et al. Fermi-level-dependent charge-to-spin current conversion by Dirac surface state of topological insulators. *Nat. Phys.* **12**, 1027–1031 (2016).
29. Rojas Sánchez, J. C. et al. Spin-to-charge conversion using Rashba coupling at the interface between non-magnetic materials. *Nat. Commun.* **4**, 2944 (2013).
30. Shen, K., Vignale, G. & Raimondi, R. Microscopic theory of the inverse Edelstein effect. *Phys. Rev. Lett.* **112**, 096601 (2014).
31. Rojas-Sánchez, J.-C. et al. Spin to charge conversion at room temperature by spin pumping into a new type of topological insulator: α -Sn films. *Phys. Rev. Lett.* **116**, 096602 (2016).
Highly efficient spin-to-charge conversion in a topological insulator is observed at room temperature.
32. Mellnik, A. R. et al. Spin-transfer torque generated by a topological insulator. *Nature* **511**, 449–451 (2014).
33. Fan, Y. et al. Magnetization switching through giant spin–orbit torque in a magnetically doped topological insulator heterostructure. *Nat. Mater.* **13**, 699–704 (2014).
34. Li, C. H. et al. Electrical detection of charge-current-induced spin polarization due to spin-momentum locking in Bi₂Se₃. *Nat. Nanotechnol.* **9**, 218–224 (2014).
Spin polarisation induced by a charge current in the surface states of a single-Dirac-cone topological insulator is conclusively demonstrated.
35. Wang, H. et al. Surface-state-dominated spin–charge current conversion in topological-insulator–ferromagnetic-insulator heterostructures. *Phys. Rev. Lett.* **117**, 076601 (2016).
36. Shiomi, Y. et al. Spin–electricity conversion induced by spin injection into topological insulators. *Phys. Rev. Lett.* **113**, 196601 (2014).
37. Nomura, A., Tashiro, T., Nakayama, H. & Ando, K. Temperature dependence of inverse Rashba–Edelstein effect at metallic interface. *Appl. Phys. Lett.* **106**, 212403 (2015).
38. Lesne, E. et al. Highly efficient and tunable spin-to-charge conversion through Rashba coupling at oxide interfaces. *Nat. Mater.* <http://www.doi.org/10.1038/nmat4726> (2016).
39. Zhang, J. et al. Band structure engineering in (Bi_{1–x}Sb_x)₂Te₃ ternary topological insulators. *Nat. Commun.* **2**, 574 (2011).
40. Wang, Z. F. et al. Creation of helical Dirac fermions by interfacing two gapped systems of ordinary fermions. *Nat. Commun.* **4**, 1384 (2013).
41. He, K. et al. Cross-over of the three-dimensional topological insulator Bi₂Se₃ to the two-dimensional limit. *Nat. Phys.* **6**, 584–588 (2010).
42. Soumyanarayanan, A. & Hoffman, J. E. Momentum-resolved STM studies of Rashba-split surface states on the topological semimetal Sb. *J. Electron Spectrosc. Relat. Phenom.* **201**, 66–73 (2015).
43. Wang, J., Chen, X., Zhu, B. & Zhang, S. Topological p–n junction. *Phys. Rev. B* **85**, 235131 (2012).
44. Castro Neto, A. H., Guinea, F., Peres, N. M. R., Novoselov, K. S. & Geim, A. K. The electronic properties of graphene. *Rev. Mod. Phys.* **81**, 109–162 (2009).
45. Castro Neto, A. H. & Guinea, F. Impurity-induced spin–orbit coupling in graphene. *Phys. Rev. Lett.* **103**, 026804 (2009).
46. Gmitra, M., Konschuh, S., Ertler, C., Ambrosch-Draxl, C. & Fabian, J. Band-structure topologies of graphene: spin–orbit coupling effects from first principles. *Phys. Rev. B* **80**, 235431 (2009).
47. Marchenko, D. et al. Giant Rashba splitting in graphene due to hybridization with gold. *Nat. Commun.* **3**, 1232 (2012).

This paper reports the first demonstration of large SOC enhancement in graphene induced by hybridization with underlying metal.

48. Rashba, E. I. Graphene with structure-induced spin-orbit coupling: spin-polarized states, spin zero modes, and quantum Hall effect. *Phys. Rev. B* **79**, 161409 (2009).
49. Ferreira, A., Rappoport, T. G., Cazalilla, M. A. & Castro Neto, A. H. Extrinsic spin Hall effect induced by resonant skew scattering in graphene. *Phys. Rev. Lett.* **112**, 066601 (2014).
50. Mendes, J. B. S. *et al.* Spin-current to charge-current conversion and magnetoresistance in a hybrid structure of graphene and yttrium iron garnet. *Phys. Rev. Lett.* **115**, 226601 (2015).
51. Balakrishnan, J., Kok Wai Koon, G., Jaiswal, M., Castro Neto, A. H. & Özyilmaz, B. Colossal enhancement of spin-orbit coupling in weakly hydrogenated graphene. *Nat. Phys.* **9**, 284–287 (2013).
52. Yankowitz, M., McKenzie, D. & LeRoy, B. J. Local spectroscopic characterization of spin and layer polarization in WSe₂. *Phys. Rev. Lett.* **115**, 136803 (2015).
53. Cheng, C. *et al.* Direct observation of spin-to-charge conversion in MoS₂ monolayer with spin pumping. Preprint at <https://arxiv.org/abs/1510.03451> (2016).
54. Wang, Z. *et al.* Strong interface-induced spin-orbit interaction in graphene on WS₂. *Nat. Commun.* **6**, 8339 (2015).
55. Vaklinova, K., Hoyer, A., Burghard, M. & Kern, K. Current-induced spin polarization in topological insulator-graphene heterostructures. *Nano Lett.* **16**, 2595–2602 (2016).
56. Bogdanov, A. & Hubert, A. Thermodynamically stable magnetic vortex states in magnetic crystals. *J. Magn. Magn. Mater.* **138**, 255–269 (1994).
57. Bogdanov, A. N. & Rößler, U. K. Chiral symmetry breaking in magnetic thin films and multilayers. *Phys. Rev. Lett.* **87**, 037203 (2001).
- This work provides the first theoretical prediction and description of magnetic skyrmions in thin films.**
58. Mühlbauer, S. *et al.* Skyrmion lattice in a chiral magnet. *Science* **323**, 915–919 (2009).
59. Yu, X.-Z. *et al.* Real-space observation of a two-dimensional skyrmion crystal. *Nature* **465**, 901–904 (2010).
60. Nagaosa, N. & Tokura, Y. Topological properties and dynamics of magnetic skyrmions. *Nat. Nanotechnol.* **8**, 899–911 (2013).
61. Heinze, S. *et al.* Spontaneous atomic-scale magnetic skyrmion lattice in two dimensions. *Nat. Phys.* **7**, 713–718 (2011).
62. Romming, N. *et al.* Writing and deleting single magnetic skyrmions. *Science* **341**, 636–639 (2013).
- This paper reports the first observation of isolated, 8-nm skyrmions in an ultrathin magnetic film, and demonstrates the creation and deletion of skyrmions.**
63. Heide, M., Bihlmayer, G. & Blügel, S. Dzyaloshinskii–Moriya interaction accounting for the orientation of magnetic domains in ultrathin films: Fe/W(110). *Phys. Rev. B* **78**, 140403 (2008).
64. Meckler, S. *et al.* Real-space observation of a right-rotating inhomogeneous cycloidal spin spiral by spin-polarized scanning tunneling microscopy in a triple axes vector magnet. *Phys. Rev. Lett.* **103**, 157201 (2009).
65. Chen, G. *et al.* Tailoring the chirality of magnetic domain walls by interface engineering. *Nat. Commun.* **4**, 2671 (2013).
66. Emori, S., Bauer, U., Ahn, S.-M., Martinez, E. & Beach, G. S. D. Current-driven dynamics of chiral ferromagnetic domain walls. *Nat. Mater.* **12**, 611–616 (2013).
67. Moreau-Lucaire, C. *et al.* Additive interfacial chiral interaction in multilayers for stabilization of small individual skyrmions at room temperature. *Nat. Nanotechnol.* **11**, 444–448 (2016).
- 30-nm skyrmions in sputtered multilayer films are observed at room temperature.**
68. Woo, S. *et al.* Observation of room-temperature magnetic skyrmions and their current-driven dynamics in ultrathin metallic ferromagnets. *Nat. Mater.* **15**, 501–506 (2016).
69. Jiang, W. *et al.* Blowing magnetic skyrmion bubbles. *Science* **349**, 283–286 (2015).
70. Fert, A., Cros, V. & Sampaio, J. Skyrmions on the track. *Nat. Nanotechnol.* **8**, 152–156 (2013).
71. Dupé, B., Hoffmann, M., Paillard, C. & Heinze, S. Tailoring magnetic skyrmions in ultra-thin transition metal films. *Nat. Commun.* **5**, 4030 (2014).
72. Yang, H., Thiaville, A., Rohart, S., Fert, A. & Chshiev, M. Anatomy of Dzyaloshinskii–Moriya interaction at Co/Pt interfaces. *Phys. Rev. Lett.* **115**, 267210 (2015).
73. Cho, J. *et al.* Thickness dependence of the interfacial Dzyaloshinskii–Moriya interaction in inversion symmetry broken systems. *Nat. Commun.* **6**, 7635 (2015).
74. Rohart, S. & Thiaville, A. Skyrmion confinement in ultrathin film nanostructures in the presence of Dzyaloshinskii–Moriya interaction. *Phys. Rev. B* **88**, 184422 (2013).
75. Sampaio, J., Cros, V., Rohart, S., Thiaville, A. & Fert, A. Nucleation, stability and current-induced motion of isolated magnetic skyrmions in nanostructures. *Nat. Nanotechnol.* **8**, 839–844 (2013).
76. Hagemester, J., Romming, N., von Bergmann, K., Vedmedenko, E. Y. & Wiesendanger, R. Stability of single skyrmionic bits. *Nat. Commun.* **6**, 8455 (2015).
77. Chen, G., Mascaraque, A., N'Diaye, A. T. & Schmid, A. K. Room temperature skyrmion ground state stabilized through interlayer exchange coupling. *Appl. Phys. Lett.* **106**, 242404 (2015).
78. Boulle, O. *et al.* Room-temperature chiral magnetic skyrmions in ultrathin magnetic nanostructures. *Nat. Nanotechnol.* **11**, 449–454 (2016).
79. Gilbert, D. A. *et al.* Realization of ground-state artificial skyrmion lattices at room temperature. *Nat. Commun.* **6**, 9462 (2015).
80. Soumyanarayanan, A. *et al.* Tunable room temperature magnetic skyrmions in Ir/Fe/Co/Pt multilayers. Preprint at <https://arxiv.org/abs/1606.06034> (2016).
81. Neubauer, A. *et al.* Topological Hall effect in the A phase of MnSi. *Phys. Rev. Lett.* **102**, 186602 (2009).
82. Schulz, T. *et al.* Emergent electrodynamics of skyrmions in a chiral magnet. *Nat. Phys.* **8**, 301–304 (2012).
83. Huang, X. & Chien, C. L. Extended skyrmion phase in epitaxial FeGe(111) thin films. *Phys. Rev. Lett.* **108**, 267201 (2012).
84. Iwasaki, J., Mochizuki, M. & Nagaosa, N. Current-induced skyrmion dynamics in constricted geometries. *Nat. Nanotechnol.* **8**, 742–747 (2013).
85. Heinonen, O., Jiang, W., Somailly, H., te Velthuis, S. G. E. & Hoffmann, A. Generation of magnetic skyrmion bubbles by inhomogeneous spin Hall currents. *Phys. Rev. B* **93**, 094407 (2016).
86. Tomasello, R. *et al.* A strategy for the design of skyrmion racetrack memories. *Sci. Rep.* **4**, 6784 (2014).
87. Kang, W. *et al.* Voltage controlled magnetic skyrmion motion for racetrack memory. *Sci. Rep.* **6**, 23164 (2016).
88. Parkin, S. S. P., Hayashi, M. & Thomas, L. Magnetic domain-wall racetrack memory. *Science* **320**, 190–194 (2008).
89. Yang, S.-H., Ryu, K.-S. & Parkin, S. Domain-wall velocities of up to 750 m s⁻¹ driven by exchange-coupling torque in synthetic antiferromagnets. *Nat. Nanotechnol.* **10**, 221–226 (2015).
90. Khvalkovskiy, A. V. *et al.* Matching domain-wall configuration and spin-orbit torques for efficient domain-wall motion. *Phys. Rev. B* **87**, 020402 (2013).
91. Moore, T. A. *et al.* High domain wall velocities induced by current in ultrathin Pt/Co/AIOx wires with perpendicular magnetic anisotropy. *Appl. Phys. Lett.* **93**, 262504 (2008).
92. Thiaville, A., Rohart, S., Jué, É., Cros, V. & Fert, A. Dynamics of Dzyaloshinskii domain walls in ultrathin magnetic films. *Europhys. Lett.* **100**, 57002 (2012).
93. Fukami, S. *et al.* Low-current perpendicular domain wall motion cell for scalable high-speed MRAM. In *2009 Symposium on VLSI Technology* 230–231 (IEEE, 2009).
94. Zhang, X., Zhou, Y., Ezawa, M., Zhao, G. P. & Zhao, W. Magnetic skyrmion transistor: skyrmion motion in a voltage-gated nanotrack. *Sci. Rep.* **5**, 11369 (2015).
95. Finocchio, G. *et al.* Skyrmion based microwave detectors and harvesting. *Appl. Phys. Lett.* **107**, 262401 (2015).
96. Manipatruni, S., Nikonov, D. E. & Young, I. A. Spin-orbit logic with magnetoelectric nodes: a scalable charge mediated nonvolatile spintronic logic. Preprint at <https://arxiv.org/abs/1512.05428> (2015).
97. Mahfouzi, F., Nikolić, B. K., Chen, S.-H. & Chang, C.-R. Microwave-driven ferromagnet-topological-insulator heterostructures: the prospect for giant spin battery effect and quantized charge pump devices. *Phys. Rev. B* **82**, 195440 (2010).
98. Cahaya, A. B., Tretiakov, O. A. & Bauer, G. E. W. Spin Seebeck power generators. *Appl. Phys. Lett.* **104**, 042402 (2014).
99. Ilan, R., De Juan, F. & Moore, J. E. Spin-based Mach–Zehnder Interferometry in topological insulator p–n junctions. *Phys. Rev. Lett.* **115**, 096802 (2015).
100. von Bergmann, K. Magnetic bubbles with a twist. *Science* **349**, 234–235 (2015).

Acknowledgements We are grateful to A.K.C. Tan and S. M. Rezende for their help preparing illustrations. We acknowledge support from the Singapore Ministry of Education (MOE), an Academic Research Fund Tier 2 (Reference No. MOE2014-T2-1-050), the National Research Foundation (NRF) of Singapore, a NRF Investigatorship (Reference No. NRF-NRFI2015-04) and the A*STAR Pharos Fund (1527400026), Singapore; and the Centre National de la Recherche Scientifique (CNRS), France, for funding this work.

Author Contributions All authors contributed equally to this work.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.F. (albert.fert@thalesgroup.com) and C.P. (christos@ntu.edu.sg).

Developmental mechanisms of stripe patterns in rodents

Ricardo Mallarino¹, Corneliu Henegar^{2,3}, Mercedes Mirasierra⁴, Marie Manceau⁵, Carsten Schradin^{6,7}, Mario Vallejo⁴, Slobodan Beronja⁸, Gregory S. Barsh^{2,3} & Hopi E. Hoekstra¹

Mammalian colour patterns are among the most recognizable characteristics found in nature and can have a profound impact on fitness. However, little is known about the mechanisms underlying the formation and subsequent evolution of these patterns. Here we show that, in the African striped mouse (*Rhabdomys pumilio*), periodic dorsal stripes result from underlying differences in melanocyte maturation, which give rise to spatial variation in hair colour. We identify the transcription factor ALX3 as a regulator of this process. In embryonic dorsal skin, patterned expression of *Alx3* precedes pigment stripes and acts to directly repress *Mitf*, a master regulator of melanocyte differentiation, thereby giving rise to light-coloured hair. Moreover, *Alx3* is upregulated in the light stripes of chipmunks, which have independently evolved a similar dorsal pattern. Our results show a previously undescribed mechanism for modulating spatial variation in hair colour and provide insights into how phenotypic novelty evolves.

The underlying mechanisms of how repetitive morphologic structures develop and evolve remain unclear. Periodic colour patterns are a useful system for the investigation of this process because of their diversity, sophistication and visual accessibility, and because the cellular, developmental and molecular mechanisms that underlie spots and stripes in mammals remain largely unknown. Traditional model organisms, such as laboratory mice (*Mus musculus*), have been instrumental for identifying genes that regulate pigment cell production, melanin synthesis and the pathways that alter the balance between two types of pigment: light pheomelanin and dark eumelanin^{1–4}. However, it is unknown to what extent these pathways explain or even contribute to the diverse array of pigment patterns seen in wild mammals. Here we take advantage of the naturally occurring coat pattern of the African striped mouse, *Rhabdomys pumilio* (Muridae), to gain insights into the processes underlying the formation and evolution of mammalian stripes, a striking and characteristic pattern that has evolved independently in many taxa, including ungulates, carnivores, rodents, marsupials, lagomorphs and primates^{3,5,6}, and is thought to confer a fitness advantage in a range of vertebrates^{7–10}.

Striped mice are diurnal, social rodents distributed throughout southwest Africa¹¹, whose dorsal coat is characterized by the presence of four dark and two light dorsal longitudinal stripes arranged in a dark–light–dark pattern (Fig. 1a). To understand how this stripe pattern is formed, we first characterized the distribution of hair and hair types across the body. In adult striped mice, hair can be classified into one of three phenotypic categories based on the individual pigment pattern—light, an unpigmented shaft with a eumelanic base; black, completely eumelanic from tip to base; and banded, a pheomelanic (yellow) shaft and eumelanic (black) base (Fig. 1b and Extended Data Fig. 1a). We quantified the proportion of each hair category in different regions along the dorsoventral axis (middle stripe, dark stripe, light stripe, flank and ventrum) and found that their proportions differed: the middle stripe and flank had similar proportions of all three hair types, but the light stripe contained mostly light hair, similar to what was found in

the ventrum, whilst the dark stripe contained mostly black hair (Fig. 1b and Extended Data Fig. 1b). Thus, variation between light and dark stripes is largely determined by changes in the distribution of unpigmented (light) and eumelanic (black) hair, rather than by differences in pigment-type switching (pheomelanin to eumelanin, and vice versa).

To investigate how such spatial differences arise, we studied skin development of striped mice between embryonic and early postnatal stages and observed stripe-like changes in both hair length and skin colour (Fig. 1c and Extended Data Fig. 2a–d). The hair length changes first became apparent at embryonic day 19 (E19; Extended Data Fig. 2a, b), but did not persist in adult mice (Extended Data Fig. 2e). By contrast, the skin colour changes were first visible at E22, creating a stable arrangement and appearance that persists as the animal grows (Fig. 1c). At birth, both phenomena are present and are correlated with both morphology and gene expression (Extended Data Fig. 2c, d, f). Thus, the patterning mechanisms that underlie adult stripes begin during embryogenesis and can be visualized by changes in the organization of mesenchymal tissues.

By postnatal day 2 (P2), light and dark stripes had similar levels of cell proliferation and hair follicle densities (Extended Data Fig. 3). However, histological sections of skin showed that hair follicles from the light stripe showed a large reduction in melanin deposition in the hair bulb relative to follicles from the dark stripe (Fig. 1d). Although hair follicles from both light and dark stripes contained pigment cells, as shown by immunohistochemistry for the pigment cell-marker KIT (Fig. 1e), those from light stripes exhibited less immunohistochemistry staining against microphthalmia-associated transcription factor (MITF) (Fig. 1f–h), a key transcription factor that promotes melanocyte differentiation and melanogenic gene expression¹². The extent of KIT staining at the base of hair follicles was indistinguishable between dark and light stripes (Fig. 1i), despite the large differences in MITF staining and visible pigmentation. We used quantitative PCR (qPCR) to measure the level of gene expression for key melanogenic genes and observed that the light stripe had lower expression of both tyrosinase (*Tyr*)

¹Howard Hughes Medical Institute, Departments of Organismic & Evolutionary Biology and Molecular & Cellular Biology, Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts 02138, USA. ²HudsonAlpha Institute for Biotechnology, Huntsville, Alabama 35806, USA. ³Department of Genetics, Stanford University School of Medicine, Stanford, California 94305, USA. ⁴Instituto de Investigaciones Biomédicas Alberto Sols (CSIC/UAM) and Cíber de Diabetes y Enfermedades Metabólicas Asociadas (Ciberdem), Madrid 28029, Spain. ⁵Center for Interdisciplinary Research in Biology, Collège de France, Paris 75005, France. ⁶Université de Strasbourg, CNRS, IPHC UMR 7178, F-67000 Strasbourg, France. ⁷School of Animal, Plant and Environmental Sciences, University of the Witwatersrand, Johannesburg 2000, South Africa. ⁸Fred Hutchinson Cancer Research Center, Seattle, Washington 98109, USA.

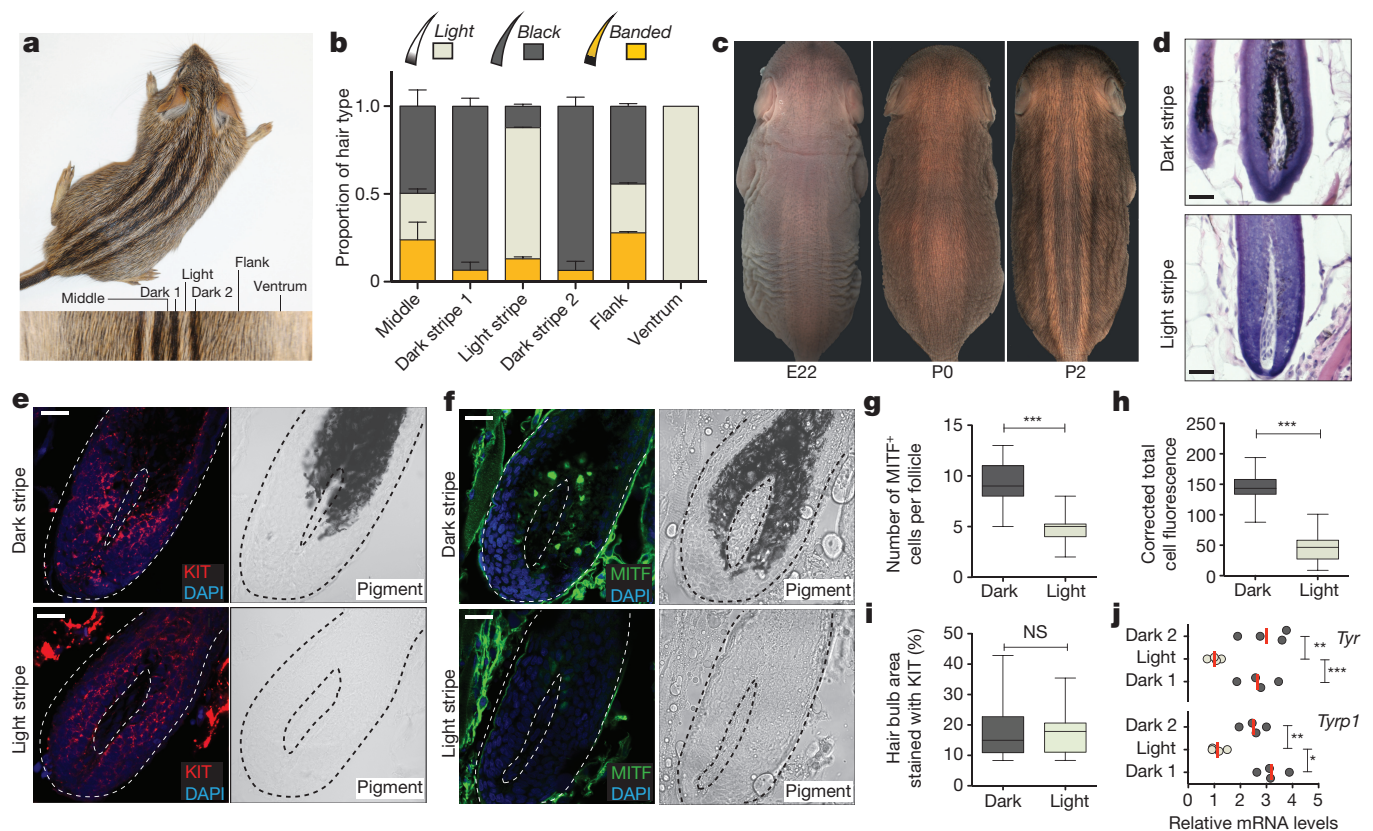


Figure 1 | Phenotypic characterization. **a**, Coat pattern of an African striped mouse. Inset, detail of the coat pattern. **b**, Proportion of each hair type in adults ($n = 5$; error bars represent s.e.m.). **c**, Dorsal pattern at embryonic day (E)22, postnatal day (P)0 and P2. **d–f**, Skin sections from P2 dark and light stripes stained with haematoxylin and eosin (**d**), and with immunohistochemistry antibodies for KIT (**e**) and MITF (**f**). Bright field images (**e**, **f**) depict pigment. Dotted lines (**e**, **f**) delineate the hair bulb. **g**, **h**, Number of detectable MITF⁺ cells (**g**, dark versus light,

two-tailed t -test; $n = 60$) and their fluorescence intensity (**h**, dark versus light, two-tailed t -test; $n = 34$). **i**, Extent of KIT stain (dark versus light, two-tailed t -test; $n = 20$, $P = 0.5429$). In **g–i**, boxes represent the 25th to 75th percentile, whiskers show the minimum and maximum value, and the horizontal line indicates the median. **j**, Quantitative PCR of melanogenic genes at P2 (ANOVA followed by a Tukey–Kramer test; $n = 4$ (individual dots); red lines depict the mean). Scale bars are 100 μm (**d**) or 50 μm (**e**, **f**). * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; NS, not significant.

and tyrosinase-related protein1 (*Tyrp1*) (Fig. 1j). Taken together, these results suggest that the differences in the proportions of light and black hair in light and dark stripes, respectively, are explained largely by differences in levels of *Mitf*, extent of melanocyte differentiation and amount of melanogenic gene expression.

Spatial differences in gene expression

We used RNA sequencing (RNA-seq) to carry out an unbiased survey of transcriptome differences that correlate with stripe identity. Three different skin regions (light stripe, dark stripe and flank) from each of three individuals per stage were investigated at four stages (E19, E22, P0 and P2; $n = 12$; 36 libraries). For E19, we used hair length as a proxy to mark and isolate incipient pigmentation stripes (Extended Data Fig. 2). Our initial analyses were carried out using, as a reference, either the *M. musculus* genome or a striped mouse transcriptome that we assembled and annotated in-house. The results from the two approaches exhibited considerable overlap but alignment to the *de novo* transcriptome reference captured a richer and more comprehensive differential signature (Extended Data Fig. 4a–c) and is described here. Of the more than 17,000 genes that we annotated in the striped mouse transcriptome, 1,062 exhibited significant differential expression between two regions (false discovery rate (FDR) < 0.1 , using the negative binomial generalized linear model implemented in DESeq2, for which both stage and region are considered factors). The largest number of differentially expressed genes between regions was observed for the flank compared to the light or the dark stripe (Extended Data Fig. 4d–f), which probably reflects a difference in body region or tissue composition rather than a colour-pattern-specific difference. Among 36 genes that were

significantly upregulated (FDR < 0.1) in the dark stripe compared to the light stripe, there is an obvious signature of melanocyte pigment production (*Tyr*, *Tyrp1*, *Mcl1*, *Oca2*, *Gpr143*, *Trpm1*; Fig. 2a and Supplementary Table 1a). Several of these genes are direct targets of *Mitf*, but we did not observe a difference in *Mitf* mRNA levels between dark- and light-striped skin, probably because *Mitf* is also expressed in non-pigmentary skin cells outside the hair follicle.

Among 28 genes that were significantly upregulated (FDR < 0.1) in the light compared to the dark stripe, a clear functional signature did not emerge; however, our attention was drawn to *Alx3*, which encodes a paired-class aristaless-like homeoprotein that has previously been implicated in fate specification of mesenchymal tissues^{13–16}. As described below, the temporal and spatial expression of *Alx3* make it a strong candidate for regulating colour differences among stripes. Furthermore, *Alx3* showed the highest fold differences in transcript abundance in the light stripe relative to dark stripe and the flank (6.73 and 4.32-fold, FDR 1.16×10^{-17} and 4.09×10^{-11} , respectively; Fig. 2a, Extended Data Fig. 4e and Supplementary Table 1a, b).

To further investigate a potential role of *Alx3* in stripe patterning, we examined its spatial and temporal expression during skin development and stripe formation. Quantitative RT–PCR showed that *Alx3* mRNA was elevated in the incipient light compared to dark stripes at E19 and that the difference increased at E22, after which it remained but was reduced at P0 (Fig. 2b). This pattern contrasts with that observed for other pigmentation-related genes including *Asip*, *Edn3*, *Tyr* and *Tyrp1* (Extended Data Fig. 5), for which differential expression is not apparent until E22 or later.

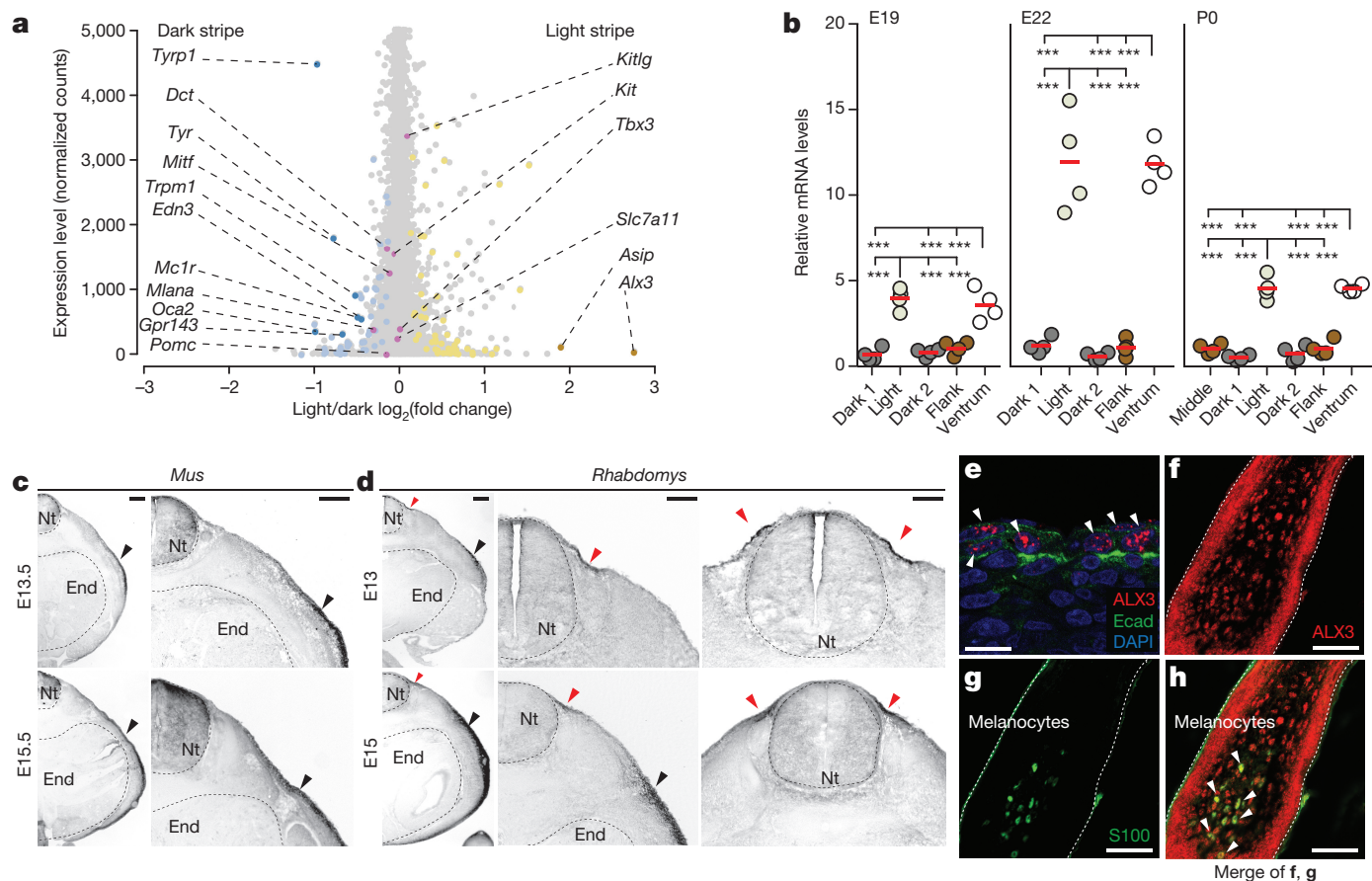


Figure 2 | *Alx3* is a candidate for regulating spatial differences in hair colour. **a**, Differential expression of transcripts in light versus dark stripes ($n = 145$; FDR < 0.1); higher expression in light (yellow) or dark stripe (blue). Known pigmentation genes are labelled, those with differential expression ($n = 9$; dark yellow, dark blue) and without differential expression ($n = 8$; pink). **b**, Quantitative PCR of *Alx3* at three stages (ANOVA followed by a Tukey–Kramer test; $n = 4$; *** $P < 0.001$. Red lines depict the mean). **c**, **d**, *In situ* hybridization shows *Alx3*

lateral and ventral mesenchyme expression (black arrowheads) in *M. musculus* (**c**) and striped mouse (**d**), and dorsal expression unique to striped mice (red arrowheads). Nt, neural tube; end, endoderm. **e–h**, Immunohistochemistry for ALX3 (arrowheads) and E-cadherin at E15 (**e**). Immunohistochemistry for ALX3 (**f**), S100 (**g**) and both merged (**h**) in P0 dorsal hair follicles. Dotted lines (**f–h**) delineate the hair bulb. Arrowheads in (**h**) show colocalization. Scale bars are 200 μ m in **c**, **d** and 50 μ m in **e–h**.

Alx3 expression during development

Using *in situ* hybridization, we examined the expression of *Alx3* during the early stages of skin development. Previous work on *Alx3* during embryogenesis (E10.5–E12.5) in *M. musculus* showed expression in neural-crest-derived mesenchyme and in lateral plate mesoderm¹³. In sections from E13.5 and E15.5 *M. musculus* embryos, we confirmed this pattern of expression, observing *Alx3* mRNA in lateral mesenchyme at E13.5, extending to and predominating in ventral mesenchyme by E15.5 (Fig. 2c). A similar pattern is present in striped mice (Fig. 2d), but striped mouse embryos show an additional domain of *Alx3* expression in the developing dorsal skin (red arrowheads in Fig. 2d), which corresponds anatomically with the future position of the stripe domain. At these early (E15) stages, ALX3 is primarily expressed in epithelial cells, as indicated by combined immunohistochemistry with E-cadherin, an epidermal compartment marker (Fig. 2e). At P0, when *Alx3* mRNA remains elevated in light compared to dark stripe skin, immunostaining for ALX3 was found in cells in developing dorsal hair bulbs (Fig. 2f). Combined immunostaining with antisera for S100 (Fig. 2g), a marker for neural crest-derived cells¹⁷, shows that hair bulb expression of ALX3 includes melanocytes as well as keratinocytes (Fig. 2h). These results demonstrate that establishment of a cellular compartment corresponding to the stripe domain occurs early during skin development, indicate that *Alx3* is a participating and/or responding factor in pattern establishment, and suggest multiple pathways—a direct effect on pigment cells and/or an indirect effect on hair bulb keratinocytes—through

which alterations in melanocyte *Mitf* expression and pattern implementation may take place.

Alx3 has a melanocyte autonomous effect

To further investigate the potential relationship between *Alx3*, *Mitf* and pigment production, we carried out both gain- and loss-of-function experiments in cultured cells. For gain-of-function experiments, a lentiviral construct in which *Alx3* and a GFP reporter were driven by a PGK promoter (LV-*Alx3*:GFP; Extended Data Fig. 6a) was transduced into B16-F1 cells, an *M. musculus* melanocyte cell line that expresses *Alx3* (Extended Data Fig. 6b); control cells were transduced with the same construct which lacked *Alx3* (LV-GFP)¹⁸ (Extended Data Fig. 6a). We found that cells carrying the full length *Alx3* construct (LV-*Alx3*:GFP) exhibited a marked decrease in mRNA levels of *Mitf* and *Silver*, a key melanogenesis gene and *Mitf* target (Extended Data Fig. 6c). For loss-of-function experiments, we used short hairpin RNAs (shRNAs) against *Alx3*. Three of the four shRNA lentiviral constructs tested (shRNA1, 2 and 3) caused *Alx3* mRNA levels to decrease relative to cells containing a scrambled shRNA control (Extended Data Fig. 6d). The same three constructs caused *Mitf* and *Silver* mRNA levels to increase. Thus, these experiments indicate that *Alx3* can negatively regulate *Mitf* in a cell-autonomous fashion.

To investigate a potential non-cell-autonomous effect of *Alx3* on *Mitf* expression, we transduced primary *M. musculus* keratinocytes with

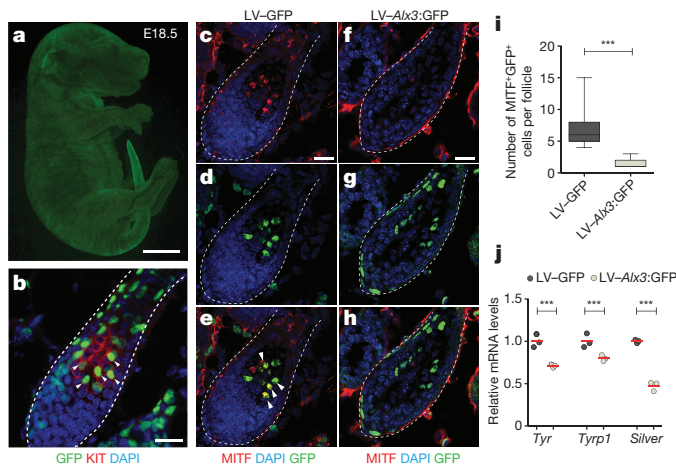


Figure 3 | *Alx3* decreases melanin synthesis *in vivo*. **a**, Ultrasound-assisted injection of a GFP-tagged lentivirus results in high-efficiency stable transduction of the skin, as shown by a *M. musculus* embryo injected at E8.5 and analysed at E18.5. **b**, A sagittal section of a hair follicle from a lentivirus-infected *M. musculus* embryo shows transduction (GFP⁺ cells) of melanocytes (GFP⁺ KIT⁺ cells indicated by arrowheads). **c–h**, Hair follicles from samples injected with LV-GFP control (**c–e**) and LV-*Alx3*:GFP (**f–h**) depicting immunohistochemistry for MITF (**c, f**), virus-transduced cells (**d, g**), and merged images showing MITF⁺ GFP⁺ cells (**e, h**) (arrowheads). Dotted lines (**b–h**) delineate the hair bulb. **i**, Number of detectable MITF⁺ cells (LV-*Alx3*:GFP versus LV-GFP; two-tailed *t*-test; *n* = 60, ****P* < 0.001). Boxes represent the 25th to 75th percentile, whiskers show the minimum and maximum value, and the horizontal line indicates the median. **j**, Quantitative PCR of mRNA levels of melanogenic genes (LV-GFP versus LV-*Alx3*:GFP; two-tailed *t*-test; *n* = 3 (individual dots); red lines depict the mean; ****P* < 0.001).

LV-*Alx3*:GFP or LV-GFP and co-cultured them with wild-type B16-F1 melanocytes using a cell-culture-insert system (Extended Data Fig. 7a). Melanocytes exposed to keratinocytes stably expressing LV-*Alx3*:GFP did not differ in their *Mitf*, *Tyr* or *Silver* mRNA levels, relative to those co-cultured with LV-GFP expressing keratinocytes (Extended Data Fig. 7b). In addition, we obtained a similar response when wild-type melanocytes were co-cultured with LV-*Alx3*:GFP transduced melanocytes (Extended Data Fig. 7c, d). These experiments show a reciprocal relationship between *Alx3* and *Mitf*, in which expression of *Mitf* mRNA and its melanogenic gene targets can be either up- or downregulated in response to inhibition or overexpression, respectively, of *Alx3* mRNA. Our results further indicate that regulation of *Mitf* expression and melanogenic gene expression by *Alx3* occurs via a melanocyte-autonomous process.

Alx3 decreases melanogenesis *in vivo*

We next examined the effect of *Alx3* *in vivo* using the same lentiviral constructs described above and ultrasound-assisted *in utero* injections into E8.5 mouse embryos. At this stage, before neural tube closure, injected lentivirus infects skin epidermis and cells originating from the neural crest, including melanocyte precursors (Fig. 3a, b). We observed GFP expression in cells in the centre of the hair follicle as well as the upper periphery; co-staining with KIT confirmed that many or most of the GFP-positive cells in the centre of the follicle were melanocytes (Fig. 3b).

Hair follicles from mice sampled at P4, a stage during which melanin synthesis is active², that were transduced with LV-*Alx3*:GFP had a marked reduction in the number of MITF⁺ GFP⁺ cells compared to the control, as revealed by immunohistochemistry (Fig. 3c–i). In addition, LV-*Alx3*:GFP-transduced melanocytes, isolated by fluorescence-activated cell sorting (FACS), showed a decrease in mRNA levels of melanin synthesis markers (*Tyrp1*, *Tyr* and *Silver*) relative to control cells (Fig. 3j). To determine whether the effect of *Alx3* overexpression reduced the number of melanocytes or just their expression of *Mitf*, we performed immunohistochemistry against SOX10, a melanocyte

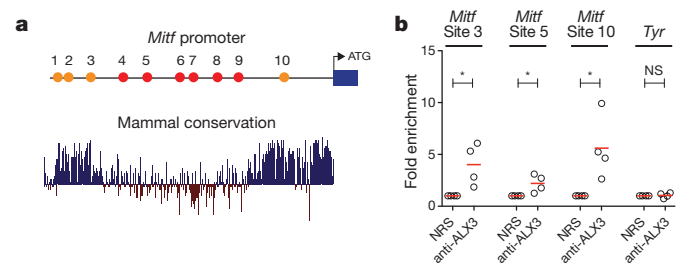


Figure 4 | *Alx3* binds to the *Mitf* promoter directly. **a**, Top, approximate locations of putative *Alx3* binding sites (labelled 1–10), conserved in *M. musculus* and striped mouse (red circles) and across mammals (orange circles), along the 1.5-kb region of the *Mitf* M promoter. ATG signals the start of the coding sequence. Bottom, plot of evolutionary conservation for the 1.5-kb region of the *Mitf* M promoter in placental mammals, identified *in silico* (<http://genome.ucsc.edu>) with sites predicted to be conserved (blue bars) and those less conserved (brown bars). **b**, Chromatin immunoprecipitation–quantitative PCR (ChIP–qPCR) assays showing amplification of *Mitf* chromatin corresponding to different regions of the promoter immunoprecipitated with an anti-ALX3 antibody or with control non-immune rabbit serum (NRS) (anti-ALX3 versus NRS; two-tailed *t*-test; *n* = 4 (individual dots); red lines depict the mean; **P* < 0.05; NS, not significant).

marker upstream of *Mitf*^{2,19–21}, and found that the number of SOX10⁺ GFP⁺ cells in LV-*Alx3*:GFP and control samples did not differ (Extended Data Fig. 8a–g). Furthermore, we did not detect a difference in the number of keratinocytes, as determined by counts of K14⁺ GFP⁺ cells inside hair follicles, or in hair follicle density (Extended Data Fig. 8h–o). The ability of elevated levels of *Alx3* to suppress expression of *Mitf* and melanocyte differentiation in experiments with *M. musculus* recapitulates the differences seen between the light and dark stripes of striped mice.

Alx3 directly represses *Mitf*

The histological and genomic data point to several pathways that contribute to stripe differences, in which an *Alx3*–*Mitf* interaction may play an early and predominant role. We next used *in vivo* and *in vitro* protein–DNA interaction experiments to gain additional insight into how *Alx3* downregulates *Mitf*. *Alx3* can selectively bind to a DNA-consensus sequence containing a TAAT motif²². The *Mitf* gene contains nine distinct promoters, of which one, 1M, is selective to melanocytes and prominently active in these cells¹². In an approximately 1.5-kb region upstream of the *M. musculus* *Mitf* M transcriptional initiation site, including regions known to be relevant for transcription factor binding²³, we identified ten candidate TAAT binding sites conserved between striped mice and *M. musculus*, four of which are also conserved across mammals (labelled 1–10; Fig. 4a and Extended Data Fig. 9). Electrophoretic mobility shift assays (EMSA) with nuclear extracts of B16-F1 cells showed that binding sites 3, 5 and 10 generated sequence-specific DNA–protein complexes (Extended Data Fig. 10a, b). The addition of ALX3 antiserum²⁴ disrupted the formation of these complexes, demonstrating that ALX3 binds to the oligonucleotide probes for these three sites (Extended Data Fig. 10a).

We then carried out chromatin immunoprecipitation (ChIP) assays on B16-F1 cells and found that *Mitf* promoter sequences that contain candidate binding sites 3, 5 and 10 were selectively amplified by qPCR from chromatin immunoprecipitated with the anti-ALX3 antiserum, but not with control non-immune serum (Fig. 4b). Promoter sequences from the *Tyr* gene, used as a control, were not amplified from the anti-ALX3 immunoprecipitated chromatin (Fig. 4b). Thus, ALX3 binds to three of the predicted binding sites immediately upstream of the melanocyte *Mitf* promoter both *in vitro* and *in vivo*.

We also tested the effect of *Alx3* overexpression on *Mitf* promoter activity in B16-F1 cells. Using a luciferase reporter cloned downstream

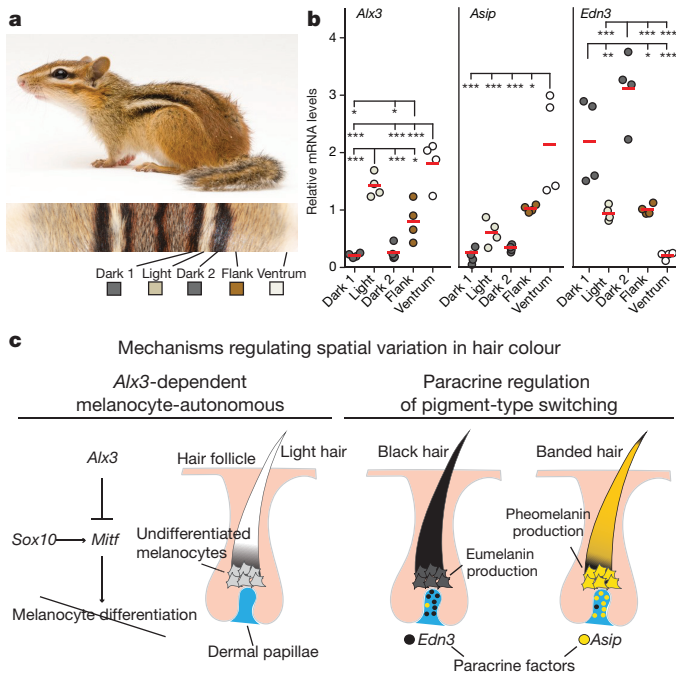


Figure 5 | Hair colour patterning mechanisms in rodents. **a**, Chipmunks independently evolved a dorsal pattern that resembles that of striped mice. Photo: J. Sartore, National Geographic Photo Ark. **b**, Quantitative PCR of *Alx3*, *Asip* and *Edn3* mRNA levels along the dorsoventral axis in chipmunks. Differences evaluated by ANOVA followed by a Tukey–Kramer test; $n = 4$ (individual dots); * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. Red lines depict the mean. **c**, Combination of a melanocyte autonomous pathway, mediated by *Alx3*, and a non-autonomous pathway, modulated by paracrine factors (*Edn3* and *Asip*), may explain variation in rodent pigmentation patterns, specifically, and in mammals, generally.

of *Mitf* promoter sequences, we found that overexpression of *Alx3* caused a decrease in luciferase expression relative to GFP overexpression (Extended Data Fig. 10c). Furthermore, when we mutated the TAAT motifs relevant for ALX3 interactions as identified by our EMSA experiments (that is, sites 3, 5 and 10), luciferase activity between GFP- and *Alx3*-expressing cells did not differ, indicating that *Alx3* could no longer suppress expression (Extended Data Fig. 10c). As a control, we mutated candidate binding site 1, which did not show binding in our EMSA experiments and found that *Alx3*-transfected cells had a decrease in luciferase activity relative to GFP-transfected cells, similar to what was seen with the wild-type *Mitf* promoter (Extended Data Fig. 10c). Together these results provide strong evidence that *Alx3* represses *Mitf* in melanocytic cells by binding directly to specific sequences in the *Mitf* promoter.

Convergent evolution of striped rodents

Colour patterns in striped mice are very similar to, but evolved independently from, those seen in Eastern chipmunks (*Tamias striatus*), a sciurid rodent that shares a last common ancestor with murids about 70 million years ago (refs 25–27) (Fig. 5a). In skin biopsies from adult Eastern chipmunks, we measured expression of *Alx3* using qPCR and found that mRNA levels were higher in the light stripe than in the dark stripe and flank (Fig. 5b). We also measured expression of *Asip* and *Edn3*, which encode paracrine regulators of pigment-type switching, and compared the results in striped mice (Extended Data Fig. 6a, b) to those in chipmunks (Fig. 5b). Overall, *Asip* mRNA levels were elevated in light areas (the light stripe, flank and ventrum in striped mice and the flank and the ventrum in chipmunks), whereas *Edn3* mRNA levels were elevated in dark areas (the dark stripes in both species). Thus, *Alx3*-mediated repression of *Mitf* in melanocytes and paracrine regulation of pigment-type switching

are common developmental mechanisms that originated repeatedly during mammalian evolution to generate stripe patterns.

Discussion

In non-mammalian vertebrates, mechanisms that underlie colour patterns depend on specialized types of pigment cells that become organized into specific arrangements during development^{3,28–30}. But in mammals, only a single type of pigment cell exists, the melanocyte, and, in general, localized differences in colour are generated by heterogeneity in gene expression rather than heterogeneity in cellular distribution^{1,3,31}. For stripes and spots in felids, pigment-type switching controlled by *Asip* and *Edn3* is the major determinant of colour pattern³², but for dorsal stripes in striped mice, suppression of melanocyte development is a main driver of colour pattern. Indeed, *Alx3*-induced inhibition of *Mitf* expression without loss of melanocytes represents a previously unrecognized evolutionary tool for the local modulation of colour pattern.

In striped mice and chipmunks, both *Alx3*-mediated repression of *Mitf* and pigment-type switching modulate hair colour. For example, increased expression of both *Asip* and *Alx3* distinguish the ventrum from the flank and increased expression of *Edn3* distinguishes the dark stripes from surrounding regions. This interplay between melanocyte-autonomous (*Alx3*-dependent) and non-autonomous (paracrine regulation of pigment-type switching) pathways may help explain patterns composed of multiple intensities of pigment (for example, pattern elements that are almost devoid of pigment (*Alx3*-mediated repression of *Mitf*) mixed with elements that are yellow or orange (pigment-type switching)) (Fig. 5c). Future efforts, directed towards developing gene-editing approaches in striped mice and chipmunks, will dissect the relative contribution of each pathway to the resulting phenotype.

From an evolutionary perspective, we show that laboratory mice, striped mice and chipmunks all express *Alx3* in ventral skin, but only the latter two species upregulate *Alx3* in their dorsum. While *Alx3* and *Asip*^{33,34} contribute to pale-coloured ventrums, we hypothesize that *Alx3* was subsequently co-opted and expressed in an additional dorsal domain, ultimately giving rise to light-coloured stripes. It is not yet clear whether the developmental cues that direct *Alx3* expression and stripe formation in striped mice and chipmunks originate from an organizing centre, such as the neuroectoderm, or arise spontaneously (for example, through a reaction-diffusion mechanism), but it is possible that the same principles operate broadly across mammalian taxa. Finally, although striped mice and chipmunks share a similar dark–light–dark stripe pattern, the location of the stripes along the dorsoventral axis differs considerably, with stripes in chipmunks situated more laterally than in striped mice. In addition, stripe number varies widely among rodents, from the single-striped grass mouse (*Lemniscomys rosalia*) to the thirteen-lined ground squirrel (*Ictidomys tridecemlineatus*). From this perspective, a detailed understanding of the spatial regulation of *Alx3* in rodents, and other striped mammals, could lead to potential insights into the evolution of gene regulation and the developmental basis of evolutionary novelty.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 4 March; accepted 16 September 2016.

Published online 2 November 2016.

- Barsh, G. S. The genetics of pigmentation: from fancy genes to complex traits. *Trends Genet.* **12**, 299–305 (1996).
- Jackson, I. J. Molecular and developmental genetics of mouse coat color. *Annu. Rev. Genet.* **28**, 189–217 (1994).
- Mills, M. G. & Patterson, L. B. Not just black and white: pigment pattern development and evolution in vertebrates. *Semin. Cell Dev. Biol.* **20**, 72–81 (2009).
- Candille, S. I. *et al.* Dorsoventral patterning of the mouse coat by *Tbx15*. *PLoS Biol.* **2**, E3 (2004).
- Morris, D. *Animal Watching. A Field Guide to Animal Behavior.* (Jonathan Cape Ltd, 1990).

6. Cloudsley-Thompson, J. L. Multiple factors in the evolution of animal coloration. *Naturwissenschaften* **86**, 123–132 (1999).
7. Caro, T., Izzo, A., Reiner, R. C. Jr, Walker, H. & Stankowich, T. The function of zebra stripes. *Nat. Commun.* **5**, 3535 (2014).
8. Brodie, E. D. III. Correlational selection for color pattern and antipredator behavior in the garter snake *Thamnophis ordinoides*. *Evolution* **46**, 1284–1298 (1992).
9. King, R. B. Color pattern polymorphism in the Lake Erie water snake, *Nerodia sipedon insularum*. *Evolution* **41**, 241–255 (1987).
10. Caro, T. The adaptive significance of coloration in mammals. *Bioscience* **55**, 125–136 (2005).
11. Schradin, C. *et al.* Social flexibility and social evolution in mammals: a case study of the African striped mouse (*Rhabdomys pumilio*). *Mol. Ecol.* **21**, 541–553 (2012).
12. Steingrimsdottir, E., Copeland, N. G. & Jenkins, N. A. Melanocytes and the microphthalmia transcription factor network. *Annu. Rev. Genet.* **38**, 365–411 (2004).
13. ten Berge, D. *et al.* Mouse *Alx3*: an aristaless-like homeobox gene expressed during embryogenesis in ectomesenchyme and lateral plate mesoderm. *Dev. Biol.* **199**, 11–25 (1998).
14. Beverdam, A., Brouwer, A., Reijnen, M., Korving, J. & Meijlink, F. Severe nasal clefting and abnormal embryonic apoptosis in *Alx3/Alx4* double mutant mice. *Development* **128**, 3975–3986 (2001).
15. Twigg, S. R. F. *et al.* Frontorhiny, a distinctive presentation of frontonasal dysplasia caused by recessive mutations in the *ALX3* homeobox gene. *Am. J. Hum. Genet.* **84**, 698–705 (2009).
16. Lakhwani, S., García-Sanz, P. & Vallejo, M. *Alx3*-deficient mice exhibit folic acid-resistant craniofacial midline and neural tube closure defects. *Dev. Biol.* **344**, 869–880 (2010).
17. Peters, E. M. J., Tobin, D. J., Botchkareva, N., Maurer, M. & Paus, R. Migration of melanoblasts into the developing murine hair follicle is accompanied by transient c-Kit expression. *J. Histochem. Cytochem.* **50**, 751–766 (2002).
18. Beronja, S., Livshits, G., Williams, S. & Fuchs, E. Rapid functional dissection of genetic networks via tissue-specific transduction and RNAi in mouse embryos. *Nat. Med.* **16**, 821–827 (2010).
19. Lee, M., Goodall, J., Verastegui, C., Ballotti, R. & Goding, C. R. Direct regulation of the Microphthalmia promoter by *Sox10* links Waardenburg-Shah syndrome (WS4)-associated hypopigmentation and deafness to WS2. *J. Biol. Chem.* **275**, 37978–37983 (2000).
20. Potterf, S. B., Furumura, M., Dunn, K. J., Arnheiter, H. & Pavan, W. J. Transcription factor hierarchy in Waardenburg syndrome: regulation of MITF expression by *SOX10* and *PAX3*. *Hum. Genet.* **107**, 1–6 (2000).
21. Elworthy, S., Lister, J. A., Carney, T. J., Raible, D. W. & Kelsh, R. N. Transcriptional regulation of *mitfa* accounts for the *sox10* requirement in zebrafish melanophore development. *Development* **130**, 2809–2818 (2003).
22. Pérez-Villamil, B., Mirasierra, M. & Vallejo, M. The homeoprotein *Alx3* contains discrete functional domains and exhibits cell-specific and selective monomeric binding and transactivation. *J. Biol. Chem.* **279**, 38062–38071 (2004).
23. Levy, C., Khaled, M. & Fisher, D. E. MITF: master regulator of melanocyte development and melanoma oncogene. *Trends Mol. Med.* **12**, 406–414 (2006).
24. García-Sanz, P., Fernández-Pérez, A. & Vallejo, M. Differential configurations involving binding of USF transcription factors and *Twist1* regulate *Alx3* promoter activity in mesenchymal and pancreatic cells. *Biochem. J.* **450**, 199–208 (2013).
25. Meredith, R. W. *et al.* Impacts of the Cretaceous Terrestrial Revolution and KPg extinction on mammal diversification. *Science* **334**, 521–524 (2011).
26. dos Reis, M. *et al.* Phylogenomic datasets provide both precision and accuracy in estimating the timescale of placental mammal phylogeny. *Proc. R. Soc. B* **279**, 3491–3500 (2012).
27. Huchon, D. *et al.* Multiple molecular evidences for a living mammalian fossil. *Proc. Natl Acad. Sci. USA* **104**, 7495–7499 (2007).
28. Rawls, J. F., Mellgren, E. M. & Johnson, S. L. How the zebrafish gets its stripes. *Dev. Biol.* **240**, 301–314 (2001).
29. Parichy, D. M. Pigment patterns: fish in stripes and spots. *Curr. Biol.* **13**, R947–R950 (2003).
30. Singh, A. P. & Nüsslein-Volhard, C. Zebrafish stripes as a model for vertebrate colour pattern formation. *Curr. Biol.* **25**, R81–R92 (2015).
31. Jackson, I. J. *et al.* Genetics and molecular biology of mouse pigmentation. *Pigment Cell Res.* **7**, 73–80 (1994).
32. Kaelin, C. B. *et al.* Specifying and sustaining pigmentation patterns in domestic and wild cats. *Science* **337**, 1536–1541 (2012).
33. Vrieling, H., Duhl, D. M., Millar, S. E., Miller, K. A. & Barsh, G. S. Differences in dorsal and ventral pigmentation result from regional expression of the mouse *agouti* gene. *Proc. Natl Acad. Sci. USA* **91**, 5667–5671 (1994).
34. Manceau, M., Domingues, V. S., Mallarino, R. & Hoekstra, H. E. The developmental role of *Agouti* in color pattern evolution. *Science* **331**, 1062–1065 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank D. Mishkind, M. Omura, J. Chupasko, P. Walsh, T. Capellini, T. Linden, K. Turner and N. Hughes for providing technical and logistical support, and C. Perdomo, A. Bendesky, C. K. Hu, D. M. Kingsley and J. M. Lassance for discussions. M.Mi. and M.V. are supported by the Spanish Ministry of Economy and Competitiveness (MINECO grants BFU2011-24245 and BFU2014-52149-R) and Instituto de Salud Carlos III. CIBERDEM is an initiative of the Instituto de Salud Carlos III. H.E.H. is an Investigator of the Howard Hughes Medical Institute.

Author Contributions R.M., M.Ma. and H.E.H. conceived the project. R.M., G.S.B. and H.E.H. designed experiments. R.M. performed cell proliferation assays, immunohistochemistry, quantitative PCR, *in situ* hybridizations, comparative sequence analysis, *in vitro* gain- and loss-of-function experiments and luciferase reporter assays. R.M. and M.Ma. collected samples and performed phenotypic characterization; R.M. and S.B. performed *in vivo* ultrasound-guided lentiviral injections. C.H. carried out the large-scale RNA experiments, including construction and annotation of the *de novo* transcriptome, and design and analysis of the RNA-seq work. M.Mi. and M.V. performed protein–DNA binding assays. C.S. provided the first embryos for pilot studies and founding members for the striped mice laboratory colony. R.M., G.S.B. and H.E.H. wrote the paper with input from all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to H.E.H. (hoekstra@oeb.harvard.edu).

Reviewer Information Nature thanks H. Arnheiter, T. Caro, M. Levine and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

Striped mouse breeding colony. F10 descendants of wild-derived striped mice (*R. pumilio*, originating from Goegap Nature Reserve, South Africa, S 29° 41.56', E 18° 1.60') were obtained from a captive colony at the University of Zurich (Switzerland) and are now maintained at Harvard University. They are kept at a 16:8 light-dark cycle and given food *ad libitum*. Developmental stages were inferred from morphological similarities with *M. musculus* embryos. Harvard University's IACUC committee approved all experiments.

Phenotypic characterization. Adults. We identified three main hair types based on their individual pigment pattern: black, banded and light. To characterize the pigment pattern along the dorsoventral axis, we quantified the proportion of each of these hair types in 1 mm hair plugs taken from each dorsal stripe, the flank and the ventrum of five adult mice. In addition, we scored the number of guard, awl and zigzag hair found in each region. To determine hair length, we placed hairs from the hair plugs on microscope slides, mounted them with glycerol and measured their length using Axiovision Microscopy Software (Zeiss).

Embryos and pups. We fixed embryos with 4% paraformaldehyde, dissected the skin and mounted it on glass slides (dermal side up). For estimating hair follicle density in pups, we detached a portion of each stripe from the muscle, while leaving the ends attached, embedded samples in OCT (Fisher Scientific), cryosectioned them transversally and stained them with haematoxylin and eosin. This technique allowed us to count individual hair follicles and assign them to the specific region to which they belonged (light stripes, dark stripes or flank). We counted the number of hair follicles and estimated the surface area of the tissue using ImageJ³⁵. Since our phenotypic characterization (Fig. 1b and Extended Data Figs 1b, 3b) and gene expression patterns determined by qPCR (Figs 1j and 2b) showed no differences between the two dark stripes, we carried out most of our analyses with dark stripe 1 (closest to the midline). We manually quantified the number of MITF⁺ cells per hair follicle, as detected with the antibody. For quantification of MITF fluorescence, we obtained images from hair follicles in the light and dark stripes, outlined stained cells, and measured the integrated density using ImageJ³⁵. To obtain the corrected total cell fluorescence (CTCF), we multiplied the area of each selected cell by the mean fluorescence of the background readings and subtracted that value from the integrated density of stained cells³⁶. To quantify the extent of KIT staining, we obtained images from hair follicles in the light and dark stripes, delineated the hair bulb area and measured the proportion of the area that was stained with KIT⁺ using ImageJ. Data were obtained from three pups and three embryos. All counts were done blind and samples were randomized. Statistical differences were determined using two-tailed *t*-tests or ANOVA (sample sizes and statistical tests used are indicated in figure legends).

Cell proliferation. We injected the peritoneum of striped mouse pregnant females with 10 µg g⁻¹ of EdU (5-ethynyl-2'-deoxyuridine) (ThermoFisher Scientific) two days before birth and collected pups at P2. To measure proliferating cells, we microdissected dark and light stripes, embedded them in OCT for posterior cryosectioning and used the Click-iT EdU Imaging kit (ThermoFisher Scientific), following the manufacturers' protocol. We counted proliferating cells in the epidermis and hair follicles from pictures of light and dark stripes. Data were obtained from three individuals. All counts were done blind and samples were randomized. Statistical differences were established using two-tailed *t*-tests (sample sizes are provided in the figure legend).

Immunohistochemistry. Striped mouse embryos were fixed in 4% paraformaldehyde, embedded in OCT/sucrose, and sectioned using a cryostat (CM 3050S, Leica). We performed immunohistochemistry using anti-MITF (Abcam 80651; 1:100), anti-ALX3 (Abcam 64985; 1:500), anti-KIT (DAKO A4502; 1:1,000), anti-E-cadherin (Millipore ECCD-2; 1:200), anti-S100 (Abcam 4066; 1:200), and anti-SOX10 (Abcam 27655; 1:100). We visualized reactions with Alexa-dye-conjugated secondary antibodies (Molecular Probes) at 1:500 dilution in 3% bovine serum albumin (BSA), PBS and Tween or with biotinylated goat anti-rabbit (Jackson Labs) and tyramide signal amplification (Perkin Elmer). For controls, we incubated sections with PBS instead of primary antibodies, but no specific cellular staining was observed. Cell nuclei were stained with DAPI (Southern Biotech). We examined sections using a LSM 700 confocal microscope and an A1 Imager (Zeiss). All pictures are representative of at least three individuals.

Quantitative PCR (qPCR). We separated the skin from the muscle and microdissected skin tissue corresponding to different regions (that is, dark stripe 1, light stripe, dark stripe 2, flank and ventrum) at the different time points indicated throughout the text. We then extracted total RNA using the fibrous tissue RNeasy kit (Qiagen), which included a DNase on-column treatment. Using qScript cDNA SuperMix (Quanta BioSciences), we generated complementary DNA (cDNA) and then performed qPCR using PerfeCTa SYBR Green FastMix (Quanta BioSciences). We used 40 cycles of amplification on an Eppendorf Mastercycler. For analysis of striped mice, we designed primers along sites that were conserved across mice and rats. For chipmunk samples, we designed primers along sites that were conserved across mice and thirteen-lined ground squirrels (all primers sequences are listed in Supplementary

Table 2). For measurements of *Mitf* expression in *M. musculus* samples, we used validated qPCR primers from the PrimerBank database³⁷. We assayed gene expression in triplicate for each sample and normalized the data using the housekeeping gene β -actin. Samples used for qPCR correspond to different individuals than those used for RNA-seq analysis. We analysed data from all qPCR experiments using the comparative *C_t* method³⁸, and established statistical significance of expression differences using either ANOVA followed by a Tukey–Kramer test or two-tailed *t*-tests (sample sizes and specific statistical tests used are given in each figure legend).

RNA sequencing. For each of the time points described in the text, we dissected skin tissue (dark stripe 1, light stripe, flank) and extracted RNA as indicated for qPCR. We used RNA from three different regions (light stripe, dark stripe and flank) from each of three individuals for four different stages (E19, E22, P0, P2; *n* = 12; 36 libraries in total). We prepared cDNA libraries for each sample using Illumina's TruSeq RNA Library Preparation Kit v2. We multiplexed individual libraries (six per lane) and sequenced them as paired-end 50-bp reads on an Illumina HiSeq 2000 instrument at the Genome Sequencing Laboratory of the HudsonAlpha Institute. We used Cutadapt software (version 1.8.1) to trim RNA-seq reads for residual adaptors and low quality sequences. Since a good quality reference genome is not currently available for the striped mouse, we used a dual exploratory strategy to assess differential expression between skin regions across various developmental stages. First, we aligned the trimmed RNA-seq reads against the *M. musculus* reference genome version using genomic sequence and transcript annotations obtained from Ensembl (release 80) and the STAR aligner software (version 2.5.0b). In parallel, we assembled trimmed RNA-seq reads into a *de novo* transcriptome using the Trinity suite of tools (version 2.1.0). The resulting *de novo* transcriptome assembly was subsequently annotated using an in-house annotation procedure supported by the human reference exon sequences retrieved from the RefSeq sequence database (GRCh37/hg19, release 55; <http://www.ncbi.nlm.nih.gov/refseq/>). Briefly, to associate a specific gene entity to each *de novo* assembled transcriptomic contig, we mapped our *de novo* assembly against a comprehensive database of reference human exon sequences using the Blast software (version 2.2.22). We retained only alignments with a significant Blast Expected Value < 10⁻⁴ for subsequent annotation purposes. Based on these alignments, we subsequently computed an *ad hoc* mapping score for each pair {assembly contig, gene entity} for which at least one significant exon alignment was identified. The mapping score was computed as the sum of the highest Blast alignment bit-scores at each position within a particular contig, associated with at least one significant alignment against an exon of the considered gene entity. Ultimately, the annotation procedure associated to each mapped contig the gene entity with the highest *ad hoc* mapping score. We then used the annotated *de novo* transcriptome assembly as reference for aligning trimmed RNA-seq reads using the Bowtie2 software (version 2.2.5).

We computed gene counts from read alignments, obtained using either the *M. musculus* reference or the *de novo* transcriptome assembly, with three software tools included in the Trinity suite: eXpress (version 1.5.1), Kallisto (0.42.4) and RSEM (version 1.2.23). We then used the individual sets of gene counts computed for each transcriptome reference and each abundance estimation tool to test for differential gene expression between samples from various skin regions with the DESeq2 package (version 1.10.1) from Bioconductor. The entire dataset (3 individuals, 3 regions, 4 stages; *n* = 36 libraries) was analysed under the DESeq2 negative binomial generalized linear model, which is a powerful and robust approach for identifying genes that are differentially expressed, either between stages or between regions. In all analyses, we used a FDR < 0.1 as a statistical significance threshold. Results presented here depict region-specific two-way comparisons across all four stages, dark versus light (Fig. 2a), light versus flank (Extended Data Fig. 4e) and dark versus flank (Extended Data Fig. 4f), in which genes assessed as significant represent the intersection between the three abundance estimation approaches implemented in the Trinity suite.

There are major changes in cell composition and skin development across the four stages we examined by RNA-seq (E19, E22, P0, P2), associated with large changes in gene expression profiles. Therefore to examine the relationship between light and dark stripes at different stages, we developed supervised learning models in which the gene expression profile at one stage was tested as a predictor of stripe identity, light versus dark, and subsequent stages. After normalization and variance stabilization using R functions implemented in DESeq2, we carried out a principal component analysis (PCA), using R functions implemented in the FactoMineR (version 1.33) package, to identify variance components associated with light versus dark phenotype across all stages. The PCA results then were used to develop supervised learning models using the R functions implemented in the randomForest package (version 4.6-12), including optimization steps based on the top 5% of the most informative gene expression profiles associated with each stage. The results demonstrate the ability of learning models based on a specific region and stage to predict region identity, light versus dark, in other stages, in which the accuracy of the models is evaluated by averaging over 30 independent iterations (Extended Data Fig. 2f).

In situ hybridizations. For *in situ* hybridization, we generated species-specific riboprobes by cloning a 545-bp fragment of *Alx3* from *M. musculus* and the striped mouse. We carried out section *in situ* hybridizations following protocols described previously^{39,40} and visualized samples using an A1 Imager (Zeiss). All pictures are representative of at least three individuals.

Cell culture experiments. We purchased B16-F1 melanoma cells from ATCC and maintained them in DMEM with 10% fetal bovine serum (FBS, Sigma-Aldrich), 100 U ml⁻¹ penicillin, and 100 µg ml⁻¹ streptomycin in a 37 °C incubator with 5% CO₂ at physiological pH 7.4. We grew cells to 70–80% confluency and performed all experiments within 10 passages. B16-F1 cells tested negative for *Mycoplasma* contamination. We cultured mouse keratinocytes and maintained them in 0.05 mM Ca²⁺ E-medium with 10% FBS serum, following previously established protocols¹⁸.

For gain-of-function experiments, we used the LV-*Alx3*:GFP and LV-GFP constructs described above. For loss-of-function experiments, we used five constructs (four specific to *Alx3* and one scrambled sequence) from existing RNAi lentiviral libraries⁴¹ (details and clone numbers listed in Supplementary Table 3). For viral infections, we plated cells in 6-well dishes at 300,000 cells per well and incubated with lentivirus in the presence of polybrene (100 µg ml⁻¹). All infections were carried out in triplicate. After two days in culture, we selected infected cells using either puromycin (2 µg ml⁻¹; shRNA constructs) or FACS (gain-of-function), and processed samples for mRNA analyses. For cell-culture-insert experiments, we plated wild-type cells on 0.4 µm transwell inserts (Falcon, BD) at 200,000 cells per ml and incubated them in plates containing a bottom layer of transduced cells (keratinocytes or melanocytes) for three days (see Extended Data Fig. 7a, c for an illustration of the experimental design).

Ultrasound-assisted *in utero* lentiviral microinjections. For construction of LV-*Alx3*:GFP, we replaced the puromycin cassette of PLKO.1, a generic lentiviral vector containing the PKG promoter⁴¹, with a fragment containing *Alx3* cDNA cloned from *M. musculus* (NM_007441.3), a P2A sequence and a histone-fluorescent protein gene fusion (Hist2h2be-eGFP). LV-GFP, which contains only the sequence coding for Hist2h2be-eGFP, was originally designed from PLKO.1 using a similar strategy¹⁸ (Addgene plasmid 25999). We carried out large-scale production of VSV-G pseudotyped lentivirus using calcium phosphate transfections of 293FT cells and helper plasmids, pMD2.G and psPAX2 (Addgene plasmids 12259 and 12260). Transfection conditions, subsequent viral concentration and titration followed established guidelines¹⁸. For injections, we anaesthetized C57Bl/6 females at E8.5 of gestation and injected embryos with 1.5 µl of a constant viral titre. We collected transduced skin from P4 pups, which we used for immunohistochemistry, following procedures outlined above, and for isolation of virus-infected melanocytes via FACS. For FACS, we used a KIT antibody (ebioscience 14-1171-81; 1:1,000), sorted KIT⁺ GFP⁺ cells directly in TRIzol LS (Invitrogen), and extracted RNA following the protocol outlined in the TRIzol LS manual. We manually quantified the number of MITF⁺ GFP⁺ and SOX10⁺ GFP⁺ cells in hair follicles, as detected with our antibodies. To determine the number of keratinocytes per follicular area, we divided the number of KIT⁺ GFP⁺ cells in hair follicles by the hair bulb area, using ImageJ³⁵. To determine hair follicle density, we obtained pictures of skin sections stained with DAPI and counted the number of hair follicles per tissue section area, using ImageJ³⁵. Data were obtained from three individuals per lentiviral construct injected. All counts were done blind and samples were randomized. Statistical differences were established using two-tailed *t*-tests (sample sizes for each analysis are indicated in figure legends).

Comparative sequence analysis. We obtained comparative sequence data from publicly available nucleotide databases at NCBI (<http://www.ncbi.nlm.nih.gov/BLAST/>). Evolutionarily conserved non-coding sequences were identified using the global sequence alignment tool incorporated in the UCSC genome browser (<http://genome.ucsc.edu>)⁴², PipMaker (<http://bio.cse.psu.edu/pipmaker>)⁴³, and LAGAN (http://lagan.stanford.edu/lagan_web/index.shtml)⁴⁴.

Electrophoretic mobility shift assays (EMSA). We conducted EMSA using nuclear extracts⁴⁵ of melanoma B16-F1 cells prepared in the presence of protease inhibitors (complete protease inhibitor cocktail; Roche) and determined protein concentrations using the Bio-Rad protein assay. Synthetic complementary oligonucleotides were annealed and labelled using [γ -³²P]-ATP and T4 kinase. We performed binding reactions at room temperature in the presence of 20,000 counts per min (c.p.m.) of radiolabelled probe (approximately 6–10 fmol) in a volume of 20 µl containing 2 µg poly (dI-dC), 20 mM HEPES (pH 7.9), 70 mM KCl, 2.5 mM MgCl₂, 1 mM dithiothreitol, 0.3 mM EDTA and 10% glycerol. We then added competitor oligonucleotides of identical (specific) or unrelated (non-specific) sequences to the probe at the indicated fold molar excess. The sequences of the oligonucleotides used are listed in Supplementary Table 2. When indicated, we added specific antiserum²⁴ or control non-immune rabbit serum (NRS) to the binding reaction. The reaction mixtures were resolved using 5% non-denaturing polyacrylamide gels, which were subsequently dried and autoradiographed.

Western blots. We prepared cell lysates from B16-F1 cells, resolved by SDS-PAGE, and blotted them onto a BioTrace PVDF membrane (Pall Corporation).

We detected ALX3 immunoreactivity with a rabbit polyclonal primary antiserum (1:4,000 dilution)²⁴ and a horseradish peroxidase-conjugated goat anti-rabbit secondary antibody (1:10,000 dilution; Bio-Rad Laboratories). To detect ACTIN we used a mouse monoclonal antibody (1:10,000 dilution, clone AC-15; Sigma) and a horseradish peroxidase-conjugated goat anti-mouse antibody (1:5,000 dilution; Bio-Rad Laboratories). We visualized immunoreactive bands using an ECL detection system (GE Healthcare).

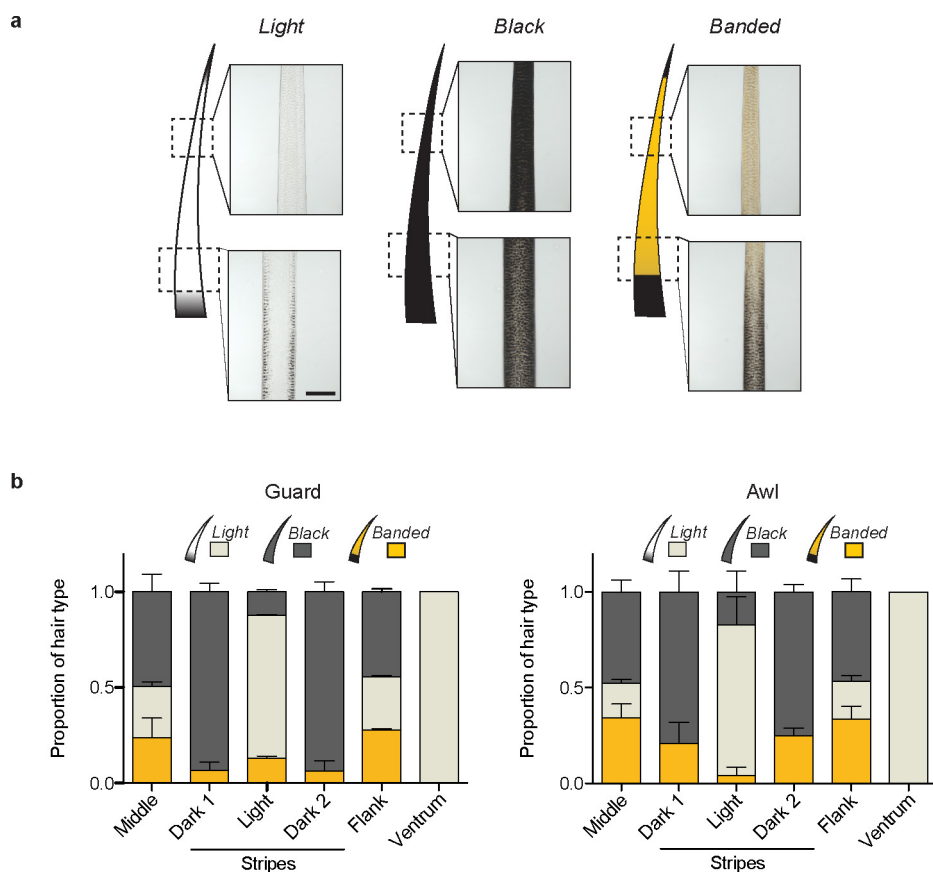
Chromatin immunoprecipitation (ChIP)–qPCR assays. We performed ChIP assays as previously described⁴⁶ using B16-F1 cells treated with 1% formaldehyde. We isolated the cross-linked protein–DNA complexes and, after sonication, we incubated chromatin with an ALX3 antiserum²⁴ or with control NRS. Next, we isolated antibody–protein–DNA complexes by incubation with protein A-sepharose. To detect bound DNA, we carried out qPCR on triplicate samples using oligonucleotide primers that amplify fragments of the *Mitf* gene corresponding to the regions containing sites 3, 5 or 10. As a control, we used promoter sequences from the *Tyr* gene as described⁴⁷. Oligonucleotides used in ChIP assays are listed in Supplementary Table 2.

Luciferase assays. We amplified a 1.5-kb region of the *Mitf* M promoter from *M. musculus* and subsequently cloned it into the SacI–HindIII sites of the pLightSwitch_Prom luciferase reporter vector (Switchgear Genomics, Active Motif). We then generated additional luciferase constructs from the wild-type construct by mutating the different *Alx3* binding motifs (TAAT to GCCG) using the Q5 site-directed mutagenesis kit (New England Biolabs). We verified all constructs by sequencing. We next transfected B16-F1 melanocytes with LV-*Alx3*:GFP and LV-GFP by using FuGENE HD (Active Motif). Using FACS, we selected the stable transfected clones and confirmed overexpression of *Alx3* by qPCR. The day before the transfection, we seeded cells at a density of 1×10^4 cells per well and 16 h later transfected them with the different *Mitf* constructs using a FuGENE HD to plasmid DNA ratio of 3:1 (300 nL FuGENE HD to 100 ng plasmid DNA per well). We then harvested cells and processed them using the LightSwitch luciferase assay kit (Switchgear Genomics) following the protocol guidelines and measured luciferase using a SpectraMax L luminometer (Molecular Devices). We normalized luciferase activity relative to luminescence from cells transfected with the pLightSwitch_Prom luciferase reporter vector (empty vector). We did not observe a difference in luciferase activity when we transfected our two stable cell lines (LV-*Alx3*:GFP and LV-GFP) with an empty vector (pLightSwitch_Prom) or a vector containing the promoter for a housekeeping gene (ACTB_PROM). We performed all luciferase experiments using five replicates per construct and established the statistical significance of luminescence differences using two-tailed *t*-tests (sample sizes for each experiment are indicated in the figure legends).

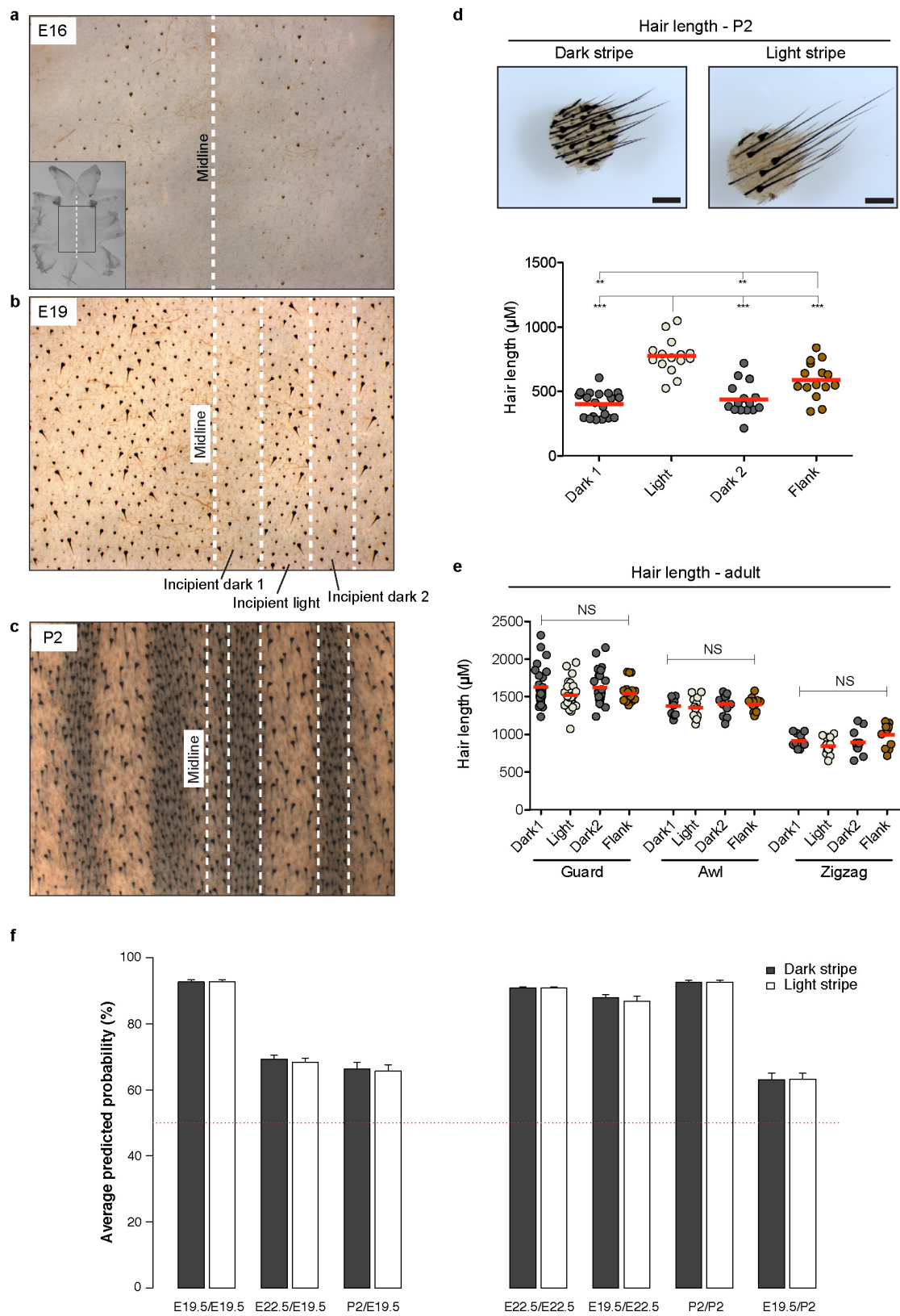
Chipmunk (*T. striatus*) samples. We collected *T. striatus* at Harvard University's Concord Field Station (Concord) using Sherman live traps (Massachusetts state permit: 027.14SCM). Chipmunks were euthanized, skin punches were taken from the different body regions, and samples were processed for qPCR as indicated above.

Data availability. The striped mouse *de novo* transcriptome assembly has been deposited in <https://datadryad.org/> under accession number doi:10.5061/dryad.7v222.

35. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).
36. McCloy, R. A. *et al.* Partial inhibition of *Cdk1* in G2 phase overrides the SAC and decouples mitotic events. *Cell Cycle* **13**, 1400–1412 (2014).
37. Spandidos, A., Wang, X., Wang, H. & Seed, B. PrimerBank: a resource of human and mouse PCR primer pairs for gene expression detection and quantification. *Nucleic Acids Res.* **38**, D792–D799 (2010).
38. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2⁻(Delta Delta C(T)) method. *Methods* **25**, 402–408 (2001).
39. Henrique, D. *et al.* Expression of a Delta homologue in prospective neurons in the chick. *Nature* **375**, 787–790 (1995).
40. Mallarino, R. *et al.* Two developmental modules establish 3D beak-shape variation in Darwin's finches. *Proc. Natl Acad. Sci. USA* **108**, 4057–4062 (2011).
41. Moffat, J. *et al.* A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell* **124**, 1283–1298 (2006).
42. Kent, W. J. *et al.* The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).
43. Schwartz, S. *et al.* PipMaker—a web server for aligning two genomic DNA sequences. *Genome Res.* **10**, 577–586 (2000).
44. Brudno, M. *et al.* LAGAN and Multi-LAGAN: efficient tools for large-scale multiple alignment of genomic DNA. *Genome Res.* **13**, 721–731 (2003).
45. Schreiber, E., Matthias, P., Müller, M. M. & Schaffner, W. Rapid detection of octamer binding proteins with 'mini-extracts', prepared from a small number of cells. *Nucleic Acids Res.* **17**, 6419 (1989).
46. Mirasiera, M. & Vallejo, M. Glucose-dependent downregulation of glucagon gene expression mediated by selective interactions between ALX3 and PAX6 in mouse alpha cells. *Diabetologia* **59**, 766–775 (2016).
47. de la Serna, I. L. *et al.* The microphthalmia-associated transcription factor requires SWI/SNF enzymes to activate melanocyte-specific genes. *J. Biol. Chem.* **281**, 20233–20241 (2006).



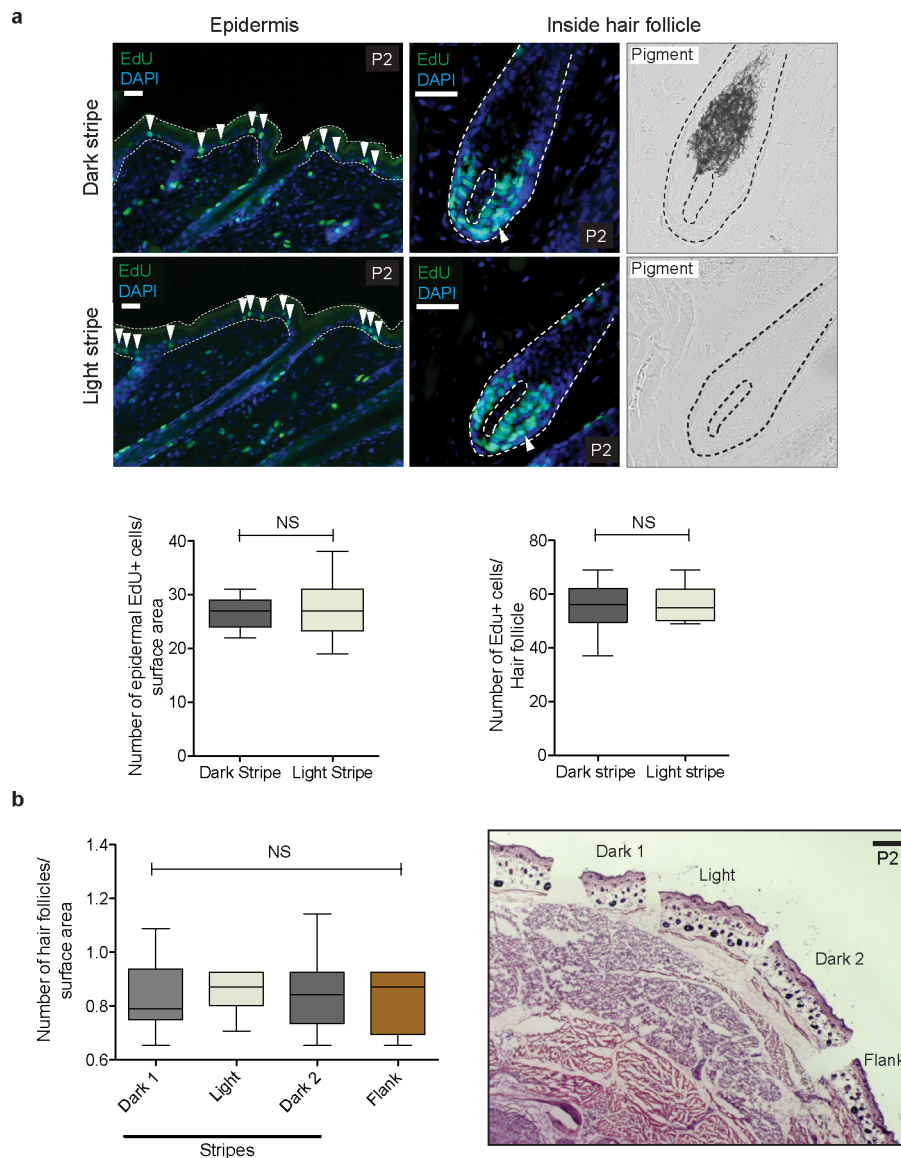
Extended Data Figure 1 | Hair characterization in adult striped mice. **a**, Striped mice have three different phenotypic categories of hair (light, black and banded) based on individual pigment pattern. All hair types have a black tip, which corresponds to structural hair features (not pigment). **b**, Relative proportion of light, black and banded guard and awl hair along the striped mouse dorsoventral axis ($n = 5$; error bars represent s.e.m.). **a**, Scale bar, 100 μm .



Extended Data Figure 2 | See next page for caption.

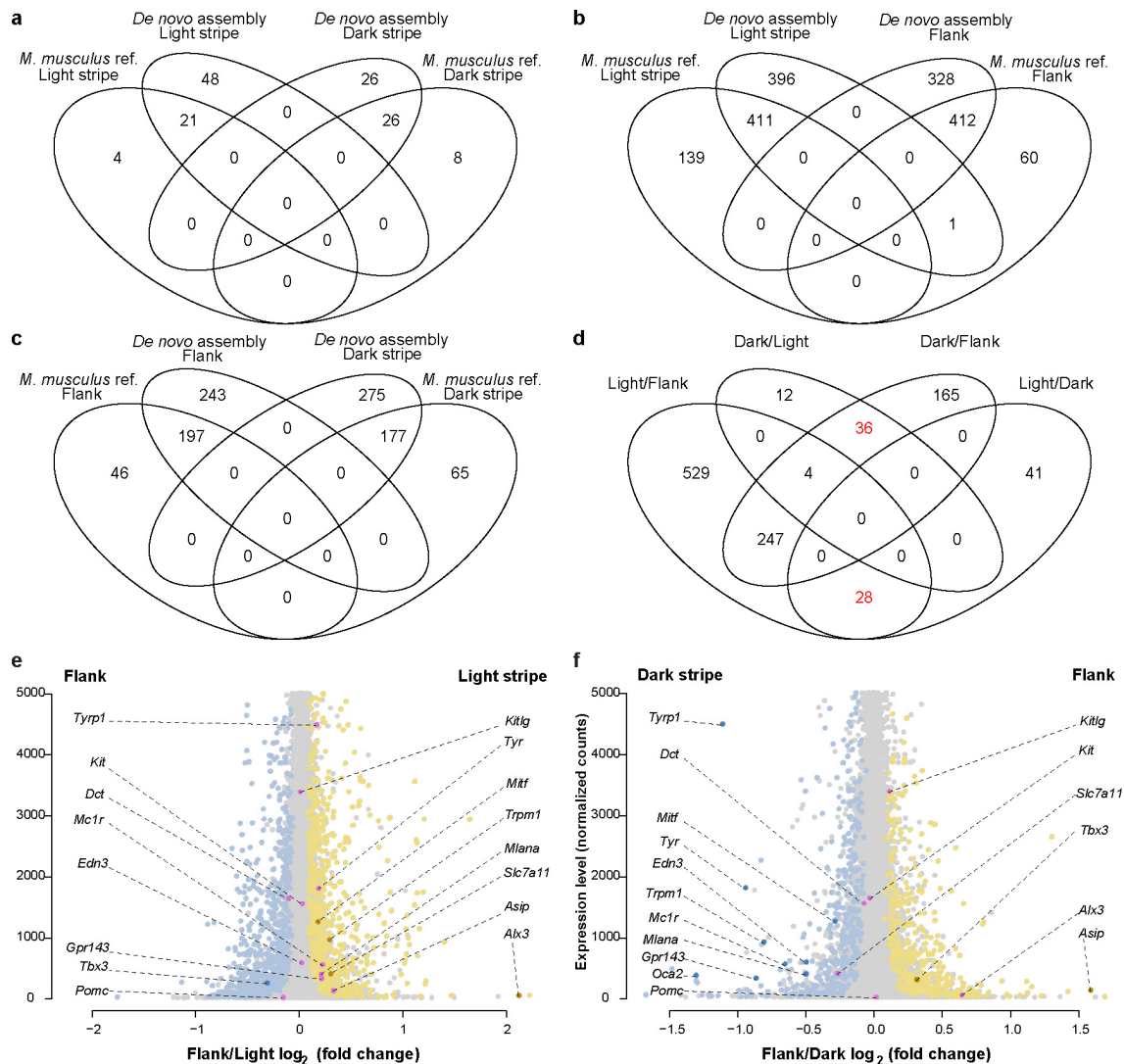
Extended Data Figure 2 | Stripe-like differences in hair length along the dorsum in striped mouse embryos and pups. a–c, Flat-mount skin preparations (dermis up) of embryos at E16 (**a**) and E19 (**b**) and P2 pups (**c**). Middle axis is indicated in all cases (midline). White dashed lines mark regions differing in hair length at E19 (**b**) and regions differing in pigmentation at P2 (**c**). Incipient pigmentation stripes are shown in (**b**). **d,** Skin punches (1 mm) and length measurements show differences between hair in the dark and light stripe of P2 individuals. Hair length differences in (**b**) (incipient stripes) correlate with those seen when pigment differences arise (**c**, **d**). Differences among dorsal regions were evaluated by ANOVA followed by a Tukey–Kramer test; $n = 15$ per region; $**P < 0.01$; $***P < 0.001$. Red lines depict the mean. **e,** Hair length measurements taken from guard, awl and zigzag hair found along the dorsum of adults. Differences among dorsal regions were evaluated

by ANOVA followed by a Tukey–Kramer test; $n = 24$ (guard), 12 (awl) and 12 (zigzag) per region; $P = 0.1736$ (guard hair), $P = 0.8006$ (awl hair), $P = 0.1038$ (zigzag hair). Red lines depict the mean. **f,** Predicted probabilities of the observed stripe-like phenotypes, as inferred by supervised learning models built and trained to recognize time-point-specific gene expression signatures of the stripes. Bars reflect average probabilities \pm s.e.m. computed from 30 consecutive iterations of the predictive model in each analysis. Labels indicate, as a ratio, the time point at which the randomForests model was applied to predict the stripe-like phenotype (the left term of the ratio), and the time point at which the model was built and trained (the right term of the ratio). The dotted line indicates the prior probability of either stripe phenotype (that is, 50% in case of only two distinct phenotypes). See Methods for details.



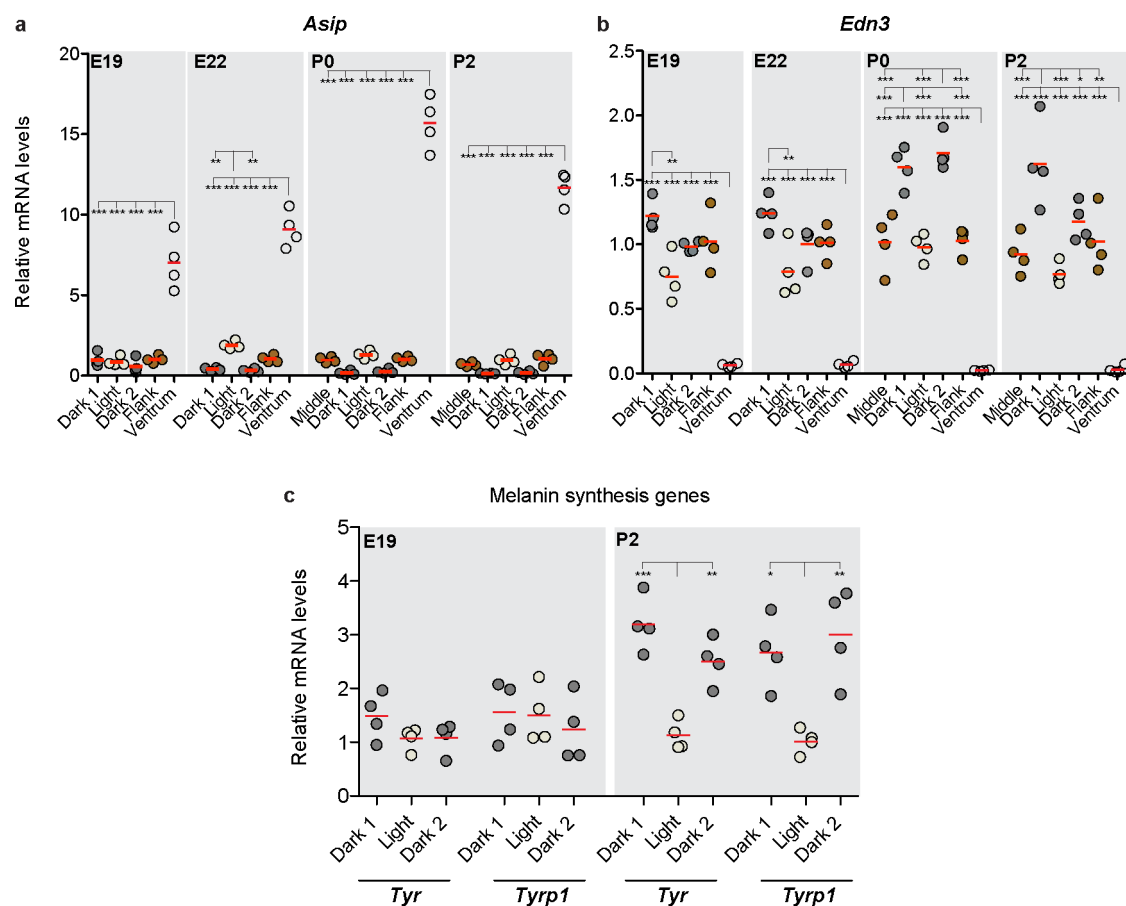
Extended Data Figure 3 | Cell proliferation and hair follicle density in P2 striped mice. a, Counts of proliferating cells, as determined by EdU labelling, in the epidermis and inside hair follicles (epidermal cells, dark stripe 1 versus light stripe, two-tailed *t*-test; $n = 15$ images per region; $P = 0.5417$; cells counted: 402 (dark stripe 1) and 444 (light stripe); intrafollicular cells, dark stripe 1 versus light stripe, two-tailed *t*-test; $n = 12$ images per region; $P = 0.7537$; cells counted: 724 (dark stripe 1) and 680 (light stripe)). Dotted lines delineate the epidermal compartment

and the hair bulb. **b,** Number of hair follicles per surface area along the dorsoventral axis. Differences among dorsal regions were evaluated by ANOVA; $n = 10$ images per region; $P = 0.4391$; number of hair follicles counted: 139 (light stripe), 141 (dark stripe 1), 132 (dark stripe 2) and 128 (flank). Bright field images in **a** depict pigment. Boxes represent the 25th to 75th percentile, whiskers show the minimum and maximum value, and the horizontal line indicates the median. Red bars indicate the mean. Scale bars in **a** 100 μm , **b** 200 μm .



Extended Data Figure 4 | RNA-seq analysis. a–d, Venn diagrams showing numbers of differentially expressed genes identified using either the *M. musculus* reference or the striped mouse *de novo* transcriptome assembly in light versus dark stripes (a), light versus flank (b), and flank versus dark (c), or differentially expressed genes in light or dark stripes versus the other skin region (light or dark stripes and the flank) (d). Genes that are specifically upregulated only in the dark or light stripes are highlighted in red. e, f, RNA-seq transcript levels (normalized gene counts) plotted as a function of differential expression (log₂ fold-change). e, The 1,777 genes demonstrating significant (FDR < 0.1) differential expression in the light stripe versus the flank are shown in yellow (higher

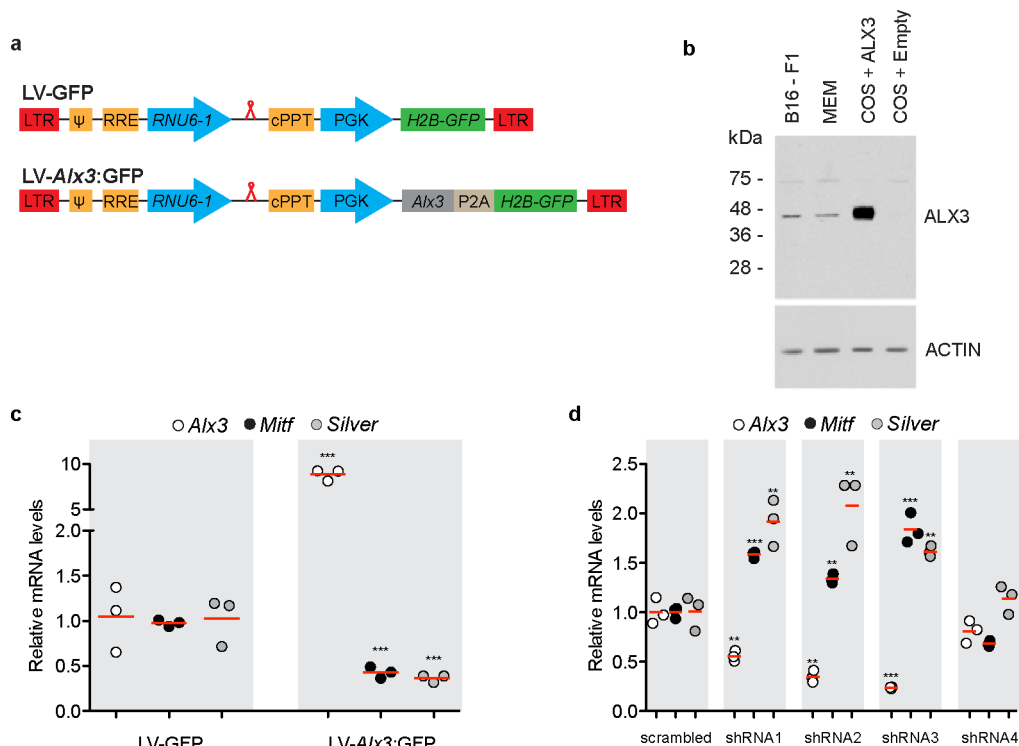
expression in the light stripe) or blue (higher expression in the flank). Four differentially expressed pigmentation-related genes are highlighted (dark yellow or dark blue), while 11 additional pigmentation-related genes that are not differentially expressed are shown in pink (Supplementary Table 1b). f, The 1,148 genes demonstrating significant (FDR < 0.1) differential expression in the flank versus the dark stripe are shown in yellow (higher expression in the flank) or blue (higher expression in the dark stripe). Eleven differentially expressed pigmentation-related genes are highlighted (dark yellow or dark blue), while six additional pigmentation-related genes that are not differentially expressed are shown in pink (Supplementary Table 1c).



Extended Data Figure 5 | Stage-specific gene expression.

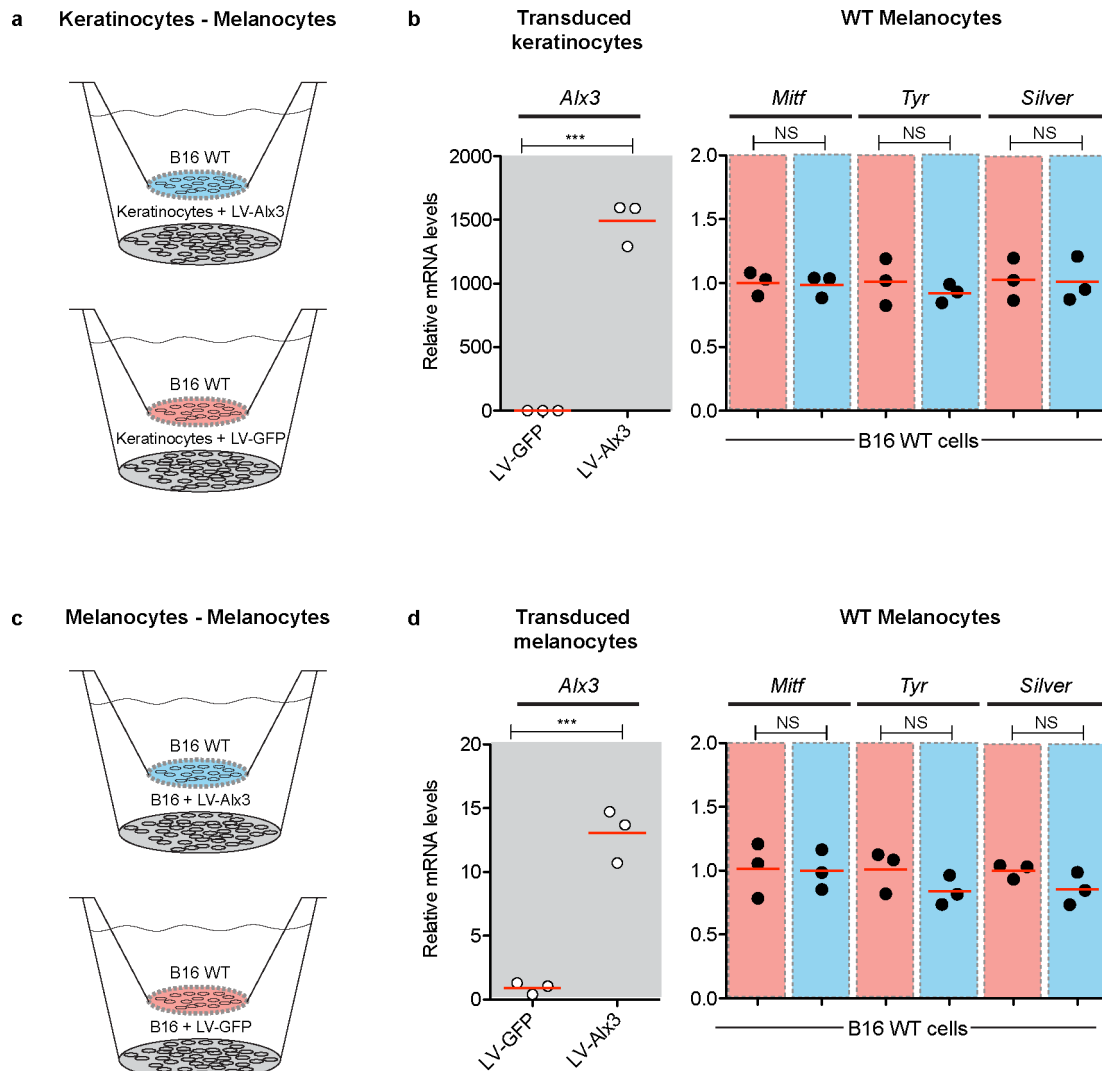
a–c, Quantitative PCR of the relative mRNA levels of pigment-type switching genes *Asip* (**a**), *Edn3* (**b**), and melanin synthesis genes *Tyr* and *Tyrp1* (**c**) in different regions of the striped mouse skin and at different

time points. Differences among stripes within each time point were evaluated by ANOVA followed by a Tukey–Kramer test; $n = 4$ per time point; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. Red bars indicate the mean.



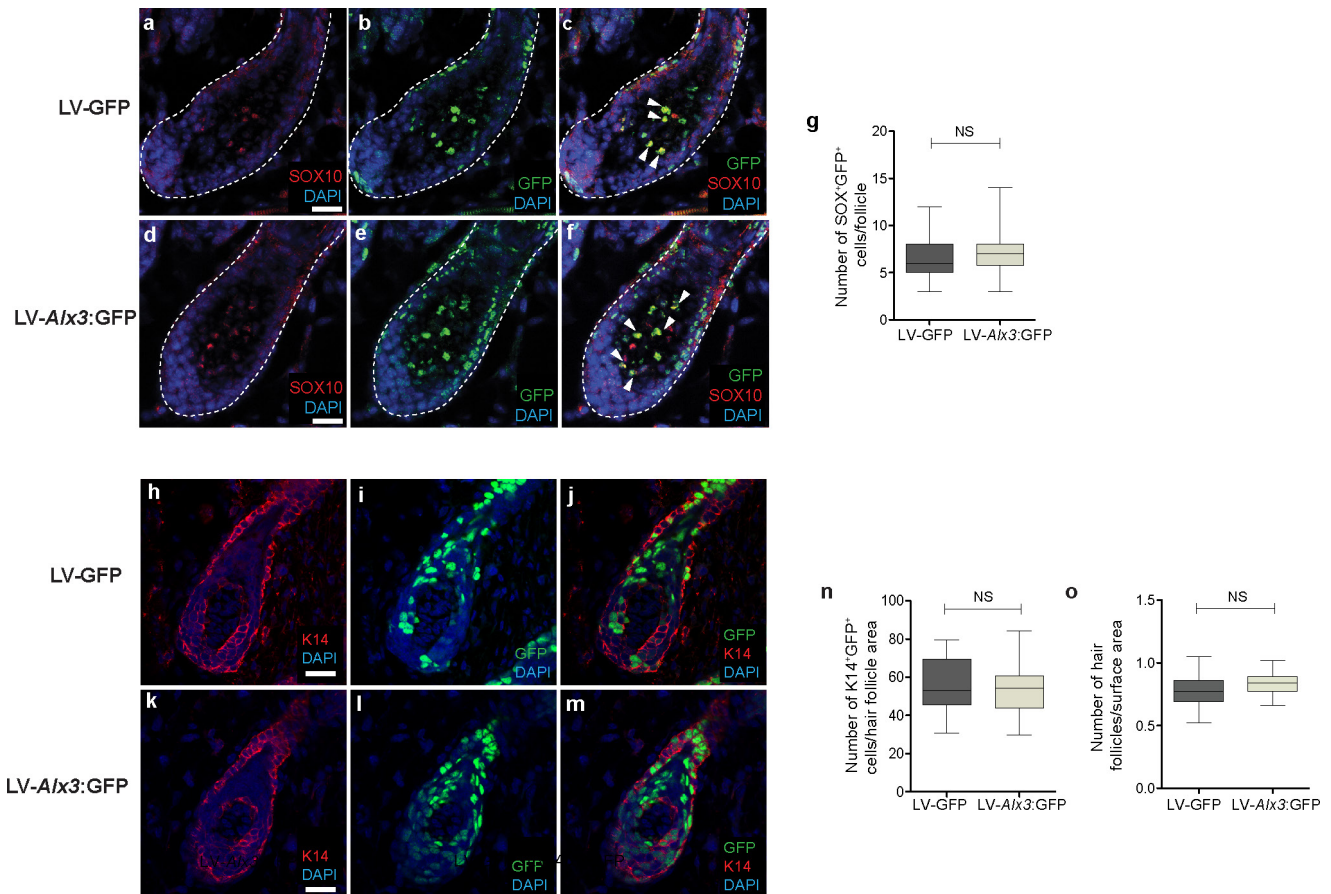
Extended Data Figure 6 | Gain- and loss-of-function experiments in cultured cells. **a**, Lentiviral constructs were modified from PLKO.1, a generic vector for expressing human RNU6-1 promoter-driven short hairpin RNAs (red loop). LTR, long terminal repeat; ψ, retroviral packaging element; RRE, Rev response element; cPPT, central polypurine tract; PGK, phosphoglycerate kinase promoter; H2B-GFP, Hist2h2be fused to GFP cDNA; P2A, 2A peptide. **b**, Western blot for expression of ALX3 in nuclear extracts of B16-F1 cells. Positive controls were extracts from mouse embryonic mesenchyme (MEM) or COS cells transfected with a pcDNA-ALX3 expression vector. COS cells transfected with an empty

pcDNA vector served as negative controls. Actin immunoreactivity is shown below for the same extracts as a control. For gel source images, see Supplementary Fig. 1. **c**, **d**, Quantitative PCR of *Alx3* (white), *Mitf* (black) and *Silver* (grey) mRNA levels in cells transduced with LV-*Alx3*:GFP (**c**), relative to cells transduced with the LV-GFP control (LV-*Alx3*:GFP versus LV-GFP, two-tailed *t*-test; *n* = 3) or shRNA lentiviral constructs (**d**), relative to cells transduced with a scrambled control (shRNA1, 2, 3 or 4 versus shRNA scrambled, two-tailed *t*-test; *n* = 3). ***P* < 0.01; ****P* < 0.001. Red bars indicate the mean.



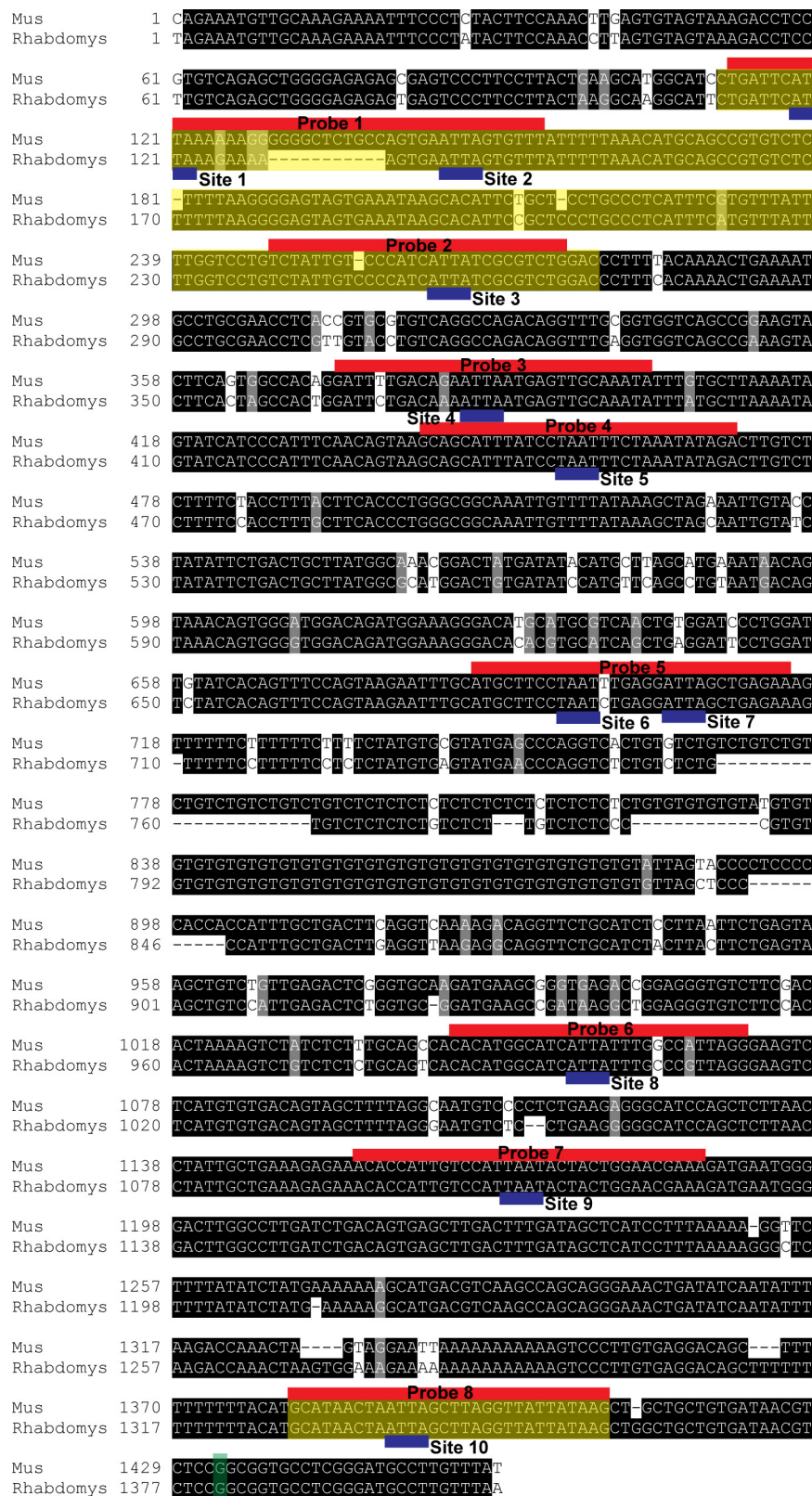
Extended Data Figure 7 | Co-culture experiments. a–d, Wild-type B16 melanocytes (B16 wild-type (WT)) were exposed to keratinocytes (**a**) or melanocytes (**c**) stably transduced with either LV-*Alx3*:GFP (LV-*Alx3* in graphs) or LV-GFP. **b, d,** Quantitative PCR of levels of *Alx3* mRNA in cells carrying the lentiviral constructs (grey panel) and of *Mitf*, *Tyr* and *Silver*

mRNA in B16 wild-type melanocytes exposed to keratinocytes (**b**) or melanocytes (**d**) transduced with LV-*Alx3*:GFP (blue panels) or LV-GFP (red panels) (LV-*Alx3*:GFP versus LV-GFP, two-tailed *t*-test; $n = 3$, *** $P < 0.001$; NS, not significant). Red bars in **b** and **d** indicate the mean.



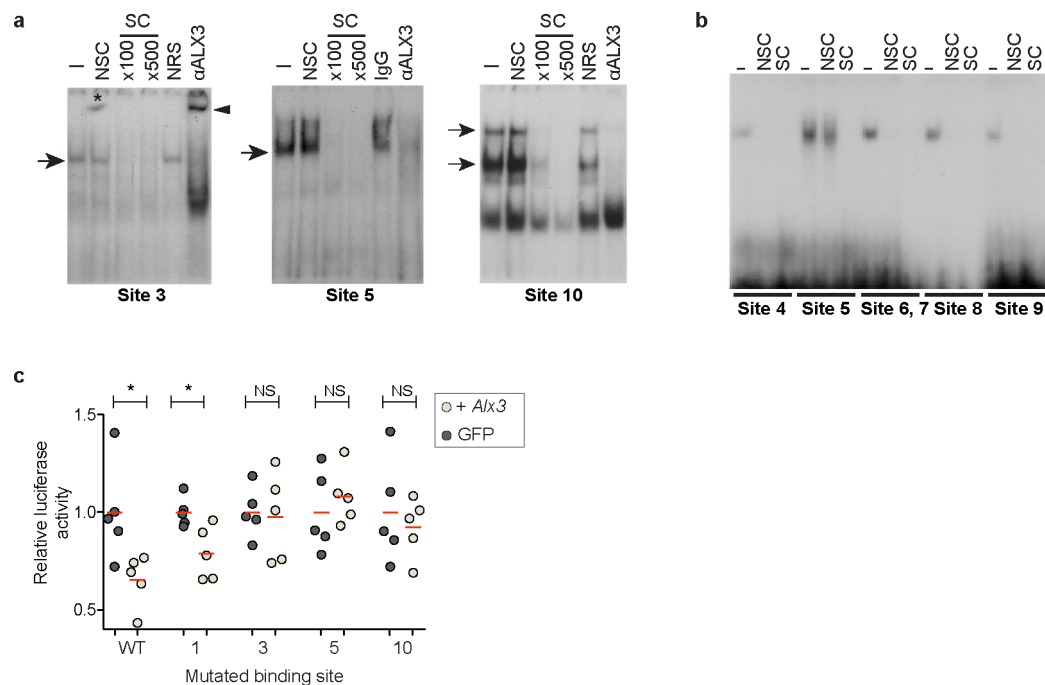
Extended Data Figure 8 | Ultrasound-guided *in utero* lentiviral injections. **a–f**, Hair follicles from embryos injected at E8.5 with lentiviruses stained for SOX10 (**a**, **d**), virus transduced cells (**b**, **e**) and merged images with arrowheads showing SOX⁺ GFP⁺ cells (**c**, **f**). Dotted lines (**a–f**) delineate the hair bulb. **g**, Number of detectable SOX10⁺ cells (LV-*Alx3*:GFP versus LV-GFP, two-tailed *t*-test; *n* = 60; *P* = 0.1173; cells counted: 426 cells (LV-*Alx3*:GFP) and 398 cells (LV-GFP)). **h–o**, Effect of *Alx3* on skin. Hair follicles from samples injected with LV-GFP control and LV-*Alx3*:GFP depicting immunohistochemistry for K14 (**h**, **k**), virus transduced cells (**i**, **l**), and merged images showing K14⁺ GFP⁺ cells (**j**, **m**).

n, Number of detectable K14⁺ GFP⁺ cells per follicular area (LV-*Alx3*:GFP versus LV-GFP, two-tailed *t*-test; *n* = 40, *P* = 0.275; average: 52.809 cells per hair follicle area (LV-*Alx3*:GFP) and 55.123 cells per hair follicle area (LV-GFP)). **o**, Hair follicle density in samples injected with viruses (LV-*Alx3*:GFP versus LV-GFP, two-tailed *t*-test; *n* = 30; *P* = 0.103; average: 0.84 hair follicles per surface area (LV-*Alx3*:GFP) and 0.794 hair follicles per surface area (LV-GFP)). Boxes represent the 25th to 75th percentile, whiskers show the minimum and maximum value, and the horizontal line indicates the median. Scale bars in **a–f** and **h–m** are 50 μM.



Extended Data Figure 9 | Alignment of a ~1.5-kb region of the *Mitf* M promoter in *M. musculus* and striped mouse. Black boxes represent conserved sequences. Mapped onto the sequences are evolutionary conserved regions of the mammalian *Mitf* M promoter identified *in silico*

(<http://genome.ucsc.edu>) (yellow), regions from which the EMSA probes were designed (red), and the transcription start site (green). The ten TAAT binding sites, conserved between *M. musculus* and striped mice (blue), which were tested are labelled sequentially.



Extended Data Figure 10 | EMSA and luciferase assays. **a**, EMSAs show the binding of nuclear proteins from B16-F1 cells to candidate-binding sites 3, 5 and 10. The absence (–) or presence of non-specific (NSC; 500-fold molar excess) or specific (SC; indicated fold molar excess) competing oligonucleotides, or the addition of ALX3 antibodies or control (NRS or IgG) is indicated. Arrows indicate complexes containing ALX3, arrowhead shows supershift for site 3 and asterisk highlights an artefact in the gel. **b**, EMSAs showing the binding of recombinant *Alx3* synthesized using a rabbit reticulocyte lysate system to the indicated sites (sites 4–9). The absence (–) or presence of non-specific (NSC) or specific (SC) competing oligonucleotides (500-fold molar excess) is indicated. Note that site 5 is the

only one showing sequence-specific binding. For gel source images, see Supplementary Fig. 1. **c**, Relative levels of luciferase activity in B16-F1 cells stably expressing GFP (dark circles) or *Alx3* (light circles) for indicated binding sites. Luciferase activity was normalized relative to cells transfected with the empty reporter vector and values are given as a fraction of luminescence for GFP-transfected cells. Differences between cells transfected with LV-*Alx3*:GFP and LV-GFP for each plasmid were evaluated using two-tailed *t*-tests; *n* = 5; **P* < 0.05; NS, not significant. Red lines depict the mean. Labels of mutated binding sites correspond to those described in Fig. 4a.

Bacteria establish an aqueous living space in plants crucial for virulence

Xiu-Fang Xin¹, Kinya Nomura¹, Kyaw Aung^{1,2}, André C. Velásquez¹, Jian Yao^{1†}, Freddy Boutrot³, Jeff H. Chang⁴, Cyril Zipfel³ & Sheng Yang He^{1,2,5,6}

High humidity has a strong influence on the development of numerous diseases affecting the above-ground parts of plants (the phyllosphere) in crop fields and natural ecosystems, but the molecular basis of this humidity effect is not understood. Previous studies have emphasized immune suppression as a key step in bacterial pathogenesis. Here we show that humidity-dependent, pathogen-driven establishment of an aqueous intercellular space (apoplast) is another important step in bacterial infection of the phyllosphere. Bacterial effectors, such as *Pseudomonas syringae* HopM1, induce establishment of the aqueous apoplast and are sufficient to transform non-pathogenic *P. syringae* strains into virulent pathogens in immunodeficient *Arabidopsis thaliana* under high humidity. *Arabidopsis* quadruple mutants simultaneously defective in a host target (AtMIN7) of HopM1 and in pattern-triggered immunity could not only be used to reconstitute the basic features of bacterial infection, but also exhibited humidity-dependent dyshomeostasis of the endophytic commensal bacterial community in the phyllosphere. These results highlight a new conceptual framework for understanding diverse phyllosphere–bacterial interactions.

The terrestrial phyllosphere represents one of the most important habitats on Earth for microbial colonization. Although the vast majority of phyllosphere microbes exhibit benign commensal associations and maintain only modest populations, adapted phyllosphere pathogens can multiply aggressively under favourable environmental conditions and cause devastating diseases. In crop fields, phyllosphere bacterial disease outbreaks typically occur after rainfall and a period of high humidity^{1–3}, consistent with the ‘disease triangle’ (host–pathogen–environment) dogma formulated more than 50 years ago⁴. The molecular basis of the effect of high humidity on bacterial infection of the phyllosphere is not understood.

Many plant and animal pathogenic bacteria, including the model phyllosphere bacterial pathogen *Pseudomonas syringae*, carry a type III secretion system (T3SS), which is used to deliver disease-promoting ‘effector’ proteins into the host cell as a primary mechanism of pathogenesis^{5,6}. Studies of how individual type III effectors promote bacterial disease in plants and animals have shown that effector-mediated suppression of host immunity is a common occurrence in both plant–bacterial^{7–9} and animal–bacterial interactions^{10,11}. However, owing to the apparent molecular complexities in bacterial diseases, which host processes must be impaired for basic bacterial pathogenesis to occur remains unknown in any plant or animal pathosystem.

Immunosuppression and pathogenesis

To investigate whether host immunity may be the only process that needs to be impaired for bacterial pathogenesis to occur in the phyllosphere, we performed infection assays in *Arabidopsis* polymutants that were severely defective in multiple immune pathways: (i) *fls2 efr cerk1* triple mutant (*fec* hereinafter), which is mutated in three major pattern recognition receptor (PRR) genes relevant to *P. syringae* pv. *tomato* strain DC3000 (*Pst*-DC3000) infection¹²; (ii) *bak1-5 bkk1-1 cerk1* triple mutant (*bbc* hereinafter; see Methods), which is compromised

in immune signalling downstream of multiple PRRs^{13,14}; and (iii) *dde2 ein2 pad4 sid2* quadruple mutant (*deps* hereinafter), which is defective in all three major defence hormone pathways (salicylic acid, jasmonate and ethylene)¹⁵. Two nonpathogenic mutant derivatives of *Pst*-DC3000 were used: the *hrcC*[−] mutant (defective in type III secretion)¹⁶ and the DC3000^{Δ28E} mutant, in which the T3SS remains intact, but 28 of 36 type III effectors (Δ28E) are deleted¹⁷. The *hrcC*[−] and DC3000^{Δ28E} mutants grew very poorly not only in wild-type *A. thaliana* Columbia-0 (Col-0) (Fig. 1a), but also in immunocompromised mutants when infiltrated into the apoplast, suggesting that host immunity is unlikely to be the only process impaired by *Pst*-DC3000 during infection.

High humidity is required for pathogenesis

During the active pathogenesis phase, phyllosphere bacterial pathogens such as *Pst*-DC3000 live mainly in the air-filled apoplast, which is connected directly to open air through epidermal pores called stomata. The water status inside the apoplast could therefore be influenced by air humidity during pathogen infection. In crop fields, phyllosphere bacterial disease outbreaks typically occur after rainfall and a period of high humidity^{1–3,18}, consistent with the ‘disease triangle’ dogma in plant pathology. In addition, one of the earliest and most common symptoms of phyllosphere bacterial diseases is the appearance of ‘water soaking’ in infected tissues, although whether water soaking has an active role in bacterial pathogenesis remains unclear. Whereas *Pst*-DC3000 multiplied to high numbers under high humidity conditions (approximately 95%; mimicking high humidity after rainfall in crop fields), it multiplied much less under low humidity conditions (<60%) (Fig. 1b), as reflected also by lower disease severity (Fig. 1c). The ability of *Pst*-DC3000 to multiply increased as humidity rose. By contrast, the *hrcC*[−] mutant multiplied poorly under all humidity conditions tested (Fig. 1d). The most aggressive infection by *Pst*-DC3000 was associ-

¹Department of Energy, Plant Research Laboratory, Michigan State University, East Lansing, Michigan 48824, USA. ²Howard Hughes Medical Institute—Gordon and Betty Moore Foundation, Michigan State University, East Lansing, Michigan 48824, USA. ³The Sainsbury Laboratory, Norwich Research Park, Norwich NR4 7UH, UK. ⁴Department of Botany and Plant Pathology and Center for Genome Research and Biocomputing, Oregon State University, Corvallis, Oregon 97331, USA. ⁵Department of Plant Biology, Michigan State University, East Lansing, Michigan 48824, USA. ⁶Plant Resilience Institute, Michigan State University, East Lansing, Michigan 48824, USA. [†]Present address: Department of Biological Sciences, Western Michigan University, Kalamazoo, Michigan 49008, USA.

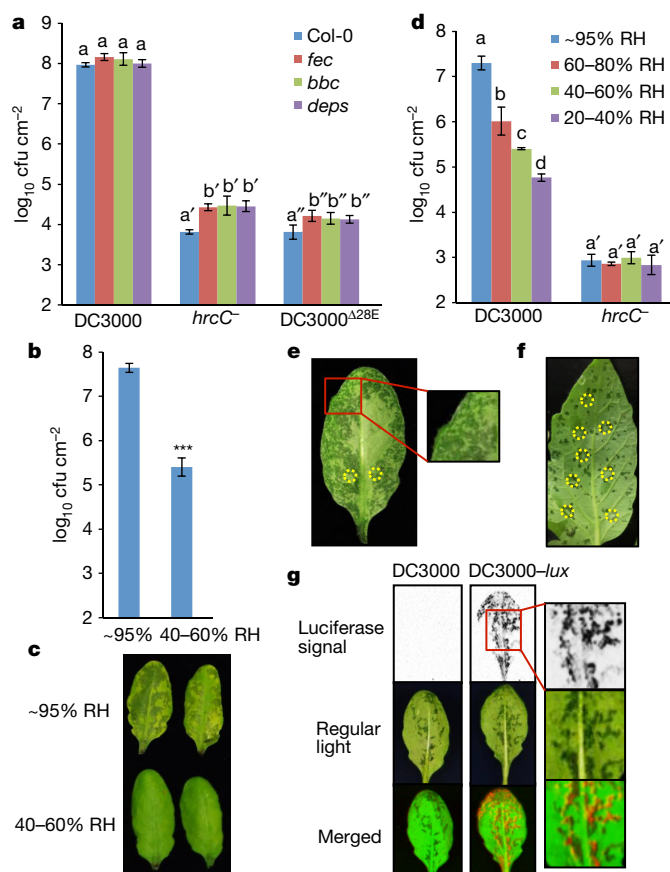


Figure 1 | Full-scale *Pst*-DC3000 infection requires high humidity and is tightly associated with apoplast 'water soaking' **a**, Bacterial populations in Col-0, *fec*, *bbc* and *deps* leaves 2 days after infiltration with bacteria at 1×10^6 cfu ml $^{-1}$. Humidity: ~95%. Two-way ANOVA with Tukey's test (significance set at $P \leq 0.05$) was performed. No significant differences were found for DC3000 populations in different plant genotypes (indicated by the same letter a), whereas differences were found for the *hrcC* $^{-}$ or DC3000 $^{\Delta 28E}$ populations in different plant genotypes, as indicated by different letters of the same type (a' versus b' for *hrcC* $^{-}$ and a'' versus b'' for DC3000 $^{\Delta 28E}$). $n = 4$ technical replicates; data are shown as mean \pm s.d. Experiments were repeated three times with similar results. **b**, **c**, Bacterial populations (**b**) and disease symptoms (**c**) 3 days after infiltration with *Pst*-DC3000 at 1×10^5 cfu ml $^{-1}$. Significant difference was determined by Student's *t*-test (two-tailed); *** $P = 1.08 \times 10^{-6}$. $n = 4$ technical replicates; data are shown as mean \pm s.d. Experiments were repeated four times with similar results. RH, relative humidity. **d**, Bacterial populations in Col-0 leaves 3 days after infiltration with bacteria at 1×10^5 cfu ml $^{-1}$. Statistical analysis as in **a**. DC3000 populations showed significant differences under different humidities (as indicated by different letters) whereas *hrcC* $^{-}$ populations did not (indicated by a'). $n = 3$ technical replicates; data are shown as mean \pm s.d. Experiments were repeated three times with similar results. **e**, The abaxial sides of Col-0 leaves 24 h after infiltration with *Pst*-DC3000 at 1×10^6 cfu ml $^{-1}$. Humidity: ~95%. Dark spots on the leaf indicate water-soaking spots. Red boxes indicate enlarged regions. **f**, A tomato leaf (cv. Castlemart) 3 days after infiltration with *Pst*-DC3000 at 1×10^4 cfu ml $^{-1}$. Humidity: ~95%. Yellow circles in **e** and **f** indicate infiltration sites. Images are representative of water-soaked leaves from more than four plants. **g**, Col-0 plants were dip-inoculated with bacteria at 2×10^8 cfu ml $^{-1}$. Humidity: ~95%. Bacterial colonies in inoculated leaves were visualized 2 days later by a charge-coupled device (top) and images of leaves were taken to show water soaking spots (middle). Bottom, merged images, with artificial red colour labelling *Pst*-DC3000-*lux* bacteria. Experiments were repeated three times. Images are representative of leaves from more than four plants.

ated with the appearance, usually within one day after infection, of water soaking in the infected *Arabidopsis* leaves under high humidity (Fig. 1e). Water-soaked spots could also be observed in *Pst*-DC3000-infected leaves of another host species, *Solanum lycopersicum* cv. Castlemart (Fig. 1f). Real-time imaging (Supplementary Video 1) showed that the initial water-soaked spots marked the areas of later disease symptoms (necrosis and chlorosis). It also showed that water soaking was a transient process, which disappeared before the onset of late disease symptoms. Using a *Pst*-DC3000 strain tagged with a luciferase reporter (DC3000-*lux*¹⁹), we found that water-soaking areas and luciferase signals were detected non-uniformly across the leaf, but they overlapped extensively (Fig. 1g and Extended Data Fig. 1), showing that water-soaked areas are where bacteria multiply aggressively in the phyllosphere before the onset of late disease symptoms.

P. syringae water-soaking effectors

The DC3000 $^{\Delta 28E}$ mutant did not cause water soaking under any experimental conditions (for example, high humidity or inoculum). We therefore transformed each of the 28 *Pst*-DC3000 effector genes back into the DC3000 $^{\Delta 28E}$ mutant individually, to identify the effector(s) that confer the ability to cause water soaking. Most effectors did not (see Fig. 2a for *avrPto*, as an example), except for *hopM1* and *avrE1*, together with their respective type III secretion chaperones *shcM* and *avrF* (Fig. 2a). This was noteworthy, because although HopM1 and AvrE1 show no sequence similarity, they were previously shown to be functionally redundant in virulence and they are highly conserved in diverse *P. syringae* strains and/or other phytopathogenic bacteria^{20,21}. Moreover, transgenic overexpression of 6 \times His:HopM1 (ref. 22) or 6 \times His:AvrE1 (ref. 23) in *Arabidopsis*, under the control of a dexamethasone-inducible promoter (10 μ M), also caused water soaking under high humidity (Fig. 2b). By contrast, transgenic expression of *AvrPto* (ref. 29), like the DC3000 $^{\Delta 28E}$ (*avrPto*) strain, did not induce water soaking. These results show that HopM1 and AvrE1, expressed either by bacteria or by transgenic overexpression in plant cells, are each sufficient to cause water soaking.

Bacterial mutant analysis showed that the *hopM1* and *avrE1* genes are necessary for *Pst*-DC3000 to cause water soaking during infection, whereas the *avrE1* $^{-}$ /*hopM1* $^{-}$ double mutant²⁰ could not cause water soaking, even when the inoculum of the *avrE1* $^{-}$ /*hopM1* $^{-}$ mutant was adjusted to reach a similar population as *Pst*-DC3000 when water soaking was assessed (Fig. 2c). By contrast, *Pst*-DC3000 and the *avrE1* $^{-}$ and *hopM1* $^{-}$ single mutants²⁰ caused strong initial water soaking (Extended Data Fig. 2a) and later disease symptoms (Extended Data Fig. 2b) and multiplied aggressively in a high-humidity-dependent manner, whereas the *avrE1* $^{-}$ /*hopM1* $^{-}$ double mutant multiplied poorly regardless of humidity conditions (Fig. 2d). Transgenic expression of 6 \times His:HopM1 in *Arabidopsis* (0.1 nM was used to induce low-level expression of HopM1 so that HopM1 alone does not cause extensive water soaking) greatly enhanced the ability of the *avrE1* $^{-}$ /*hopM1* $^{-}$ double mutant to cause water soaking and multiply extensively under high humidity conditions (Extended Data Fig. 2c, d). These results showed that, unlike the other 34 effectors present in the *avrE1* $^{-}$ /*hopM1* $^{-}$ double mutant, the full virulence functions of HopM1 and AvrE1 are uniquely dependent on external high humidity.

Next, we investigated why the virulence functions of HopM1 and AvrE1 were dependent on external humidity. We hypothesized that the primary function of HopM1 and AvrE1 is for creating an aqueous apoplast (that is, bacteria 'prefer' to live in an aqueous environment in the apoplast), the maintenance of which requires high humidity as the leaf apoplast is directly connected to open air through stomata. If so, it may be possible to substitute the function of HopM1 and AvrE1 by simply providing water to the apoplast. To test this hypothesis directly, we performed transient water supplementation experiments in which Col-0 plants infiltrated with the *avrE1* $^{-}$ /*hopM1* $^{-}$ mutant were transiently kept water-soaked, for the first 12–16 h to mimic the kinetics of transient water soaking that normally occur during *Pst*-DC3000

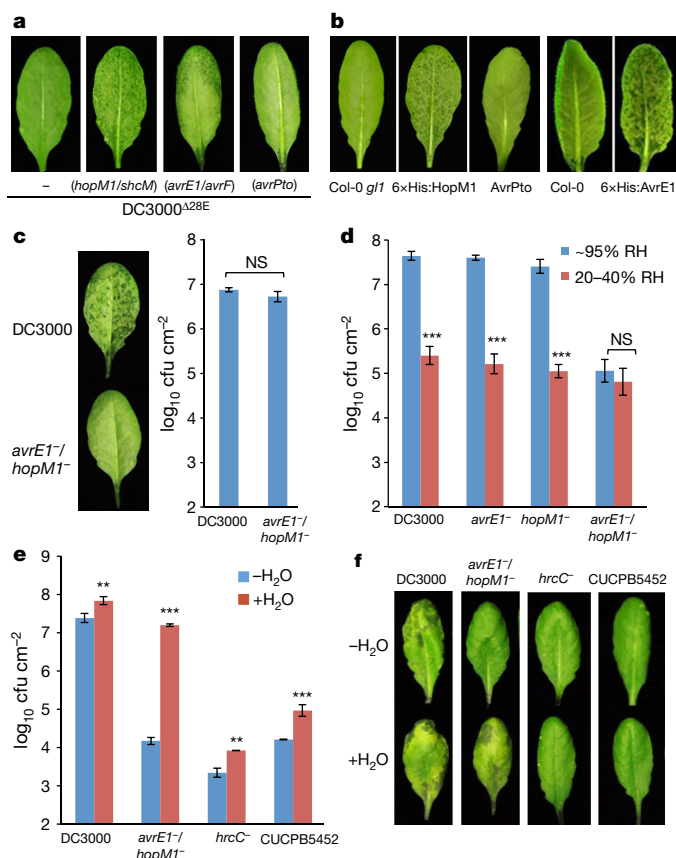


Figure 2 | Type III effectors AvrE1 and HopM1 are necessary and sufficient to cause water soaking. **a**, Col-0 leaves 24 h after infiltration with bacteria ($1-2 \times 10^8$ cfu ml $^{-1}$). Humidity: ~95%. **b**, Leaves of transgenic 6xHis:HopM1 (ref. 22), 6xHis:AvrE1 (ref. 23) or AvrPto (ref. 29) plants after spraying with 10 μ M dexamethasone (DEX; to induce effector gene expression). Humidity: ~95%. Col-0 or Col-0 *gl1* (*glabra1*) plants were non-transgenic parental controls. Images are representative of leaves from more than four plants. **c**, Col-0 leaves (left) and bacterial populations (right) 24 h after infiltration with *Pst*-DC3000 (1×10^6 cfu ml $^{-1}$) or the *avrE1* $^{-}$ /*hopM1* $^{-}$ strain (1×10^7 cfu ml $^{-1}$). Humidity: ~95%. Student's *t*-test (two-tailed); NS, not significant ($P = 0.104$). $n = 6$ technical replicates from three independent experiments ($n = 2$ in each experiment); data are shown as mean \pm s.d. **d**, Bacterial populations in Col-0 plants 3 days after infiltration with bacteria at 2×10^5 cfu ml $^{-1}$. *** $P = 1.07 \times 10^{-6}$, 8.07×10^{-7} and 5.95×10^{-7} for DC3000, the *avrE1* $^{-}$ mutant and the *hopM1* $^{-}$ mutant, respectively, for differences in bacterial population between different humidities, as determined by Student's *t*-test (two-tailed); NS, not significant ($P = 0.13$). $n = 4$ technical replicates; data are shown as mean \pm s.d. Experiments were repeated three times. **e**, **f**, Bacterial populations (**e**) and images (**f**) of Col-0 leaves 3 days after infiltration with bacteria at 1×10^5 cfu ml $^{-1}$. In the '-H₂O' treatment, plants were air-dried normally (for about 2 h) and then kept under high humidity (approximately 95%). In the '+H₂O' treatment, plants were kept under high (80–95%) humidity after syringe-infiltration to allow slow evaporation of water (for about 16 h, until no visible apoplast water was seen). ** $P = 8.29 \times 10^{-3}$ and 1.14×10^{-3} for DC3000 and *hrcC* $^{-}$, respectively and *** $P = 7.61 \times 10^{-7}$ and 9.82×10^{-4} for *avrE1* $^{-}$ /*hopM1* $^{-}$ and CUCPB5452, respectively, indicate significant differences between -H₂O and +H₂O treatments as determined by Student's *t*-test (two-tailed). $n = 3$ technical replicates; data are shown as mean \pm s.d. Experiments were repeated three times.

infection (Supplementary Video 1). Transient apoplast water supplementation was sufficient to restore the multiplication (100- to 1,000-fold) of the *avrE1* $^{-}$ /*hopM1* $^{-}$ mutant almost to the level of *Pst*-DC3000 (Fig. 2e), as well as to induce appearance of severe disease symptoms (Fig. 2f). By contrast, the controls, *Pst*-DC3000, the *hrcC* $^{-}$ mutant and CUCPB5452 (which expresses the *avrE1* and *hopM1* genes but has

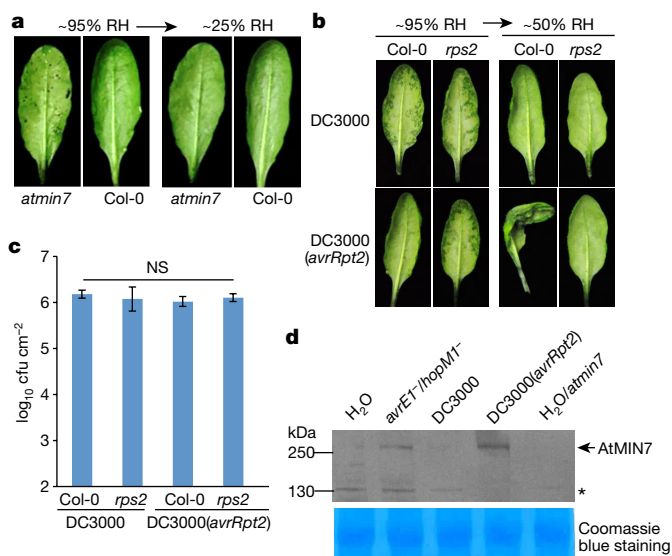


Figure 3 | Effects of AtMIN7 and effector-triggered immunity on water soaking. **a**, *atmin7* leaves, but not Col-0 leaves, showed partial water soaking 48 h after dip-inoculation with the *avrE1* $^{-}$ /*hopM1* $^{-}$ mutant at 1×10^8 cfu ml $^{-1}$. Humidity: ~95%. Water soaking disappeared after transition to low humidity (~25%) to allow evaporation of apoplast water. Images are representative of leaves from more than four plants. **b**, **c**, ETI prevents apoplast water soaking. Col-0 and *rps2* leaves were infiltrated with *Pst*-DC3000 (1×10^6 cfu ml $^{-1}$) or *Pst*-DC3000(*avrRpt2*) (1×10^7 cfu ml $^{-1}$ for Col-0 and 1×10^6 cfu ml $^{-1}$ for *rps2* plants). Plants were kept under high humidity (~95%) for 24 h to observe water soaking and then shifted to low humidity (~50%) for 4 h to observe ETI-associated tissue collapse. Images were taken before and after low humidity exposure (**b**) and bacterial populations were determined 24 h after infiltration to show similar population levels (**c**). Statistical analysis of data in **c** was performed by one-way ANOVA with Tukey's test (significance set at $P \leq 0.05$), and no significant difference was detected. $n = 3$ technical replicates; data are shown as mean \pm s.d. Experiments were repeated three times. **d**, AtMIN7 protein is stabilized during ETI as revealed by immunoblot. Col-0 or *atmin7* leaves were infiltrated with bacteria (1×10^7 cfu ml $^{-1}$; ref. 25) or H₂O and kept under high humidity (~95%) for 24 h before protein extraction. Asterisk indicates a non-specific band. Coomassie blue staining shows equal loading. See Supplementary Fig. 1 for uncropped images.

much reduced virulence owing to deletion of other type III effectors²⁴) grew only slightly better (<10 fold) with transient water supplementation (Fig. 2e). These results demonstrate that the primary virulence function of HopM1 and AvrE1 can be effectively substituted by transiently supplying water to the apoplast.

The host target of HopM1 for modulating water soaking

To investigate the mechanism by which HopM1 creates the aqueous apoplast, we focused on the host targets of HopM1 in *Arabidopsis*. We have previously shown that HopM1 is targeted to the trans-Golgi-network-early endosome (TGN/EE) in the host cell and mediates proteasome-dependent degradation of several host proteins, including AtMIN7, which is a TGN/EE-localized ADP ribosylation factor–guanine nucleotide exchange factor involved in vesicle trafficking^{22,25,26}. Although the *atmin7* mutant plant showed partially increased bacterial multiplication^{22,25}, the exact role of AtMIN7 during pathogen infection remains unknown. A previous study showed that the virulence function of HopM1 is fundamentally different from that of canonical immunosuppressing effectors, such as AvrPto¹⁷. As our data showed a role for HopM1 in creating water soaking, we tested the possibility that AtMIN7 may modulate apoplast water soaking in response to bacterial infection. Notably, we found that apoplast water soaking occurred in the *atmin7* mutant plant in the absence of HopM1/AvrE1 (that is, during infection by the *avrE1* $^{-}$ /*hopM1* $^{-}$ mutant; Fig. 3a and Extended

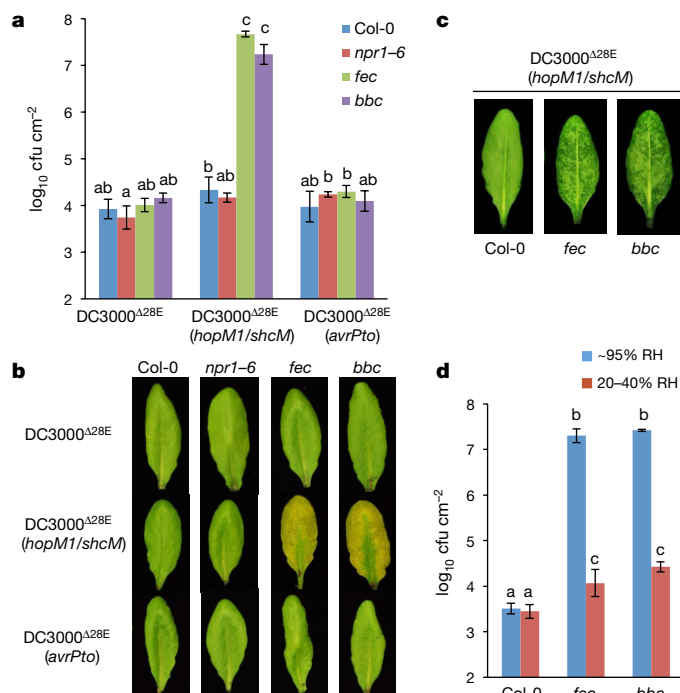


Figure 4 | *hopM1/shcM* transform the non-pathogenic DC3000 $\Delta 28E$ mutant into a highly virulent pathogen in PTI-deficient mutant plants in a humidity-dependent manner. **a**, **b**, Bacterial populations (**a**) and disease symptoms (**b**) 3 days after infiltration with the indicated bacteria at 1×10^6 cfu ml^{-1} . Humidity: ~95%. Statistical analysis was performed by one-way ANOVA with Tukey's test (significance set at $P \leq 0.05$). Bacterial populations indicated by different letters (a, b and c) are significantly different (ab is not significantly different from a or b). $n = 4$ technical replicates; data are shown as mean \pm s.d. Experiments were repeated three times. **c**, Water-soaking symptoms were recorded 24 h after inoculation. **d**, Bacterial populations 3 days after infiltration with DC3000 $\Delta 28E$ (*hopM1/shcM*) at 1×10^6 cfu ml^{-1} under indicated humidities. Statistical analysis as in **a**. Bacterial populations indicated by different letters (a, b and c) are significantly different. $n = 4$ technical replicates; data are shown as mean \pm s.d. Experiments were repeated three times. Images are representative of leaves from at least four plants.

Data Fig. 3c), and the *avrE1*[−]/*hopM1*[−] mutant could multiply in *atmin7* mutant plant (Extended Data Fig. 3a, b). Thus, genetic removal of AtMIN7 is sufficient to mimic the virulence function of HopM1, albeit partially, in causing apoplast water soaking. The *atmin7* mutant plant is defective in endocytic recycling of plasma membrane proteins and has an abnormal plasma membrane²⁶, suggesting that HopM1 degrades AtMIN7, probably to compromise host plasma membrane integrity as a mechanism to create an infection-promoting aqueous apoplast (Extended Data Fig. 4).

If apoplast water soaking is an essential step in pathogenesis, we hypothesized that plants may have evolved defence mechanisms to counter it. Indeed, we found that effector-triggered immunity (ETI)²⁷ induced by *Pst*-DC3000(*avrRpt2*) blocked water soaking, even when the inoculum of *Pst*-DC3000(*avrRpt2*) was increased to reach a population similar to *Pst*-DC3000 when water soaking was assessed (Fig. 3b, c and Extended Data Fig. 5a, b). When transferred from high (approximately 95%) to low (around 50%) humidity, *Pst*-DC3000(*avrRpt2*)-infected leaves quickly wilted, indicating extensive ETI-associated programmed cell death. By contrast, *Pst*-DC3000-infected, water-soaked leaves returned to pre-infection healthy appearance (Fig. 3b), indicating little host cell death during apoplast water soaking. Furthermore, *Pst*-DC3000(*avrRpt2*)-triggered ETI stabilized the AtMIN7 protein²⁵ (Fig. 3d). These results therefore show a previously unknown interaction between bacterial virulence (creating apoplast

water soaking) and host defence (preventing apoplast water soaking), which is linked in part to AtMIN7 stability, leading to modulation of apoplast water availability.

Reconstitution of *P. syringae* infection

As apoplast water soaking seems to be a key process in bacterial pathogenesis, we investigated a model in which suppression of pattern-triggered immunity (PTI) and creation of apoplast water soaking are two principal pathogenic processes sufficient for bacterial infection of the phyllosphere to occur. We infected Col-0 and two PTI-compromised triple mutant plants (*fec* and *bbc*) with DC3000 $\Delta 28E$, DC3000 $\Delta 28E$ (*avrPto*) or DC3000 $\Delta 28E$ (*hopM1/shcM*) and found that DC3000 $\Delta 28E$ (*hopM1/shcM*), but not DC3000 $\Delta 28E$ or DC3000 $\Delta 28E$ (*avrPto*), caused strong water soaking, multiplied aggressively (almost to the level of *Pst*-DC3000) and produced prominent disease symptoms in the *fec* and *bbc* mutant plants (Fig. 4a–c) in a high humidity-dependent manner (Fig. 4d). Furthermore, unlike PTI mutants, the *npr1-6* mutant plant, which is defective in salicylic-acid-dependent defence (Extended Data Fig. 6a–c), did not show greatly increased DC3000 $\Delta 28E$ (*hopM1/shcM*) multiplication (Fig. 4a). Thus, a combination of defective PTI and the presence of an aqueous-apoplast-inducing effector (HopM1) could almost fully convert a non-pathogenic mutant into a virulent pathogen in the *Arabidopsis* phyllosphere.

If immunosuppression and creation of apoplast water soaking are two principal pathogenic processes sufficient for bacterial infection of the phyllosphere, we hypothesized that it might be possible to construct a multi-host-target mutant that simulates the two processes. This mutant plant might allow an otherwise non-pathogenic mutant bacterium (for example, the *hrcC*[−] mutant) to colonize the phyllosphere, thereby reconstituting basic features of a phyllosphere bacterial infection. For this purpose, we mutated the *AtMIN7* gene in PTI mutants (*fec* and *bbc*) and generated *atmin7 fls2 efr cerk1* (*mfec* hereinafter) and *atmin7 bak1-5 bkk1-1 cerk1* (*mbbc* hereinafter) quadruple mutants using CRISPR–Cas9 technology (see Methods and Extended Data Fig. 7a). The quadruple mutant plants display a similar morphology to wild-type Col-0 plants (Extended Data Fig. 7b) and have a tendency to show some water-soaking spots, especially in mature leaves, under high humidity (Extended Data Fig. 7c, d). Notably, these mutants allowed the non-pathogenic *hrcC*[−] mutant to multiply aggressively under high (approximately 95%) humidity, to a final population that was about 100-fold higher than in Col-0 plants 5 days after inoculation, with the *mbbc* plants showing a greater susceptibility than the *mfec* plants (Fig. 5a). In addition, in these quadruple mutant plants, the *hrcC*[−] mutant induced prominent disease chlorosis and necrosis (Fig. 5b and Extended Data Fig. 7e), which were not observed for the *hrcC*[−] strain in Col-0, *atmin7* or PTI mutants. Thus, dual disruption of AtMIN7 and PTI signalling is sufficient to generate the basic features of a model phyllosphere bacterial disease. Consistent with these data, transient water supplementation to the leaf apoplast was sufficient to enhance the growth of the *hrcC*[−] mutant in the *bbc* triple mutant, but not in Col-0 plants (Fig. 5c). To our knowledge, this is the first infectious model disease, in plant or animal, for which basic pathogenesis has been generated using biologically relevant host-target mutants.

Dyshomeostasis of commensal bacteria

The inability of the non-pathogenic *hrcC*[−] mutant to multiply aggressively in the wild-type phyllosphere resembles that of the commensal bacterial community that resides in the apoplast of healthy leaves. Consistent with this, only low levels of the endophytic phyllosphere bacterial community were detectable in wild-type Col-0 plants (Fig. 5d). However, after plants were shifted from regular growth conditions (around 60% relative humidity, day 0; Fig. 5d) to high humidity conditions (approximately 95% relative humidity), the *mfec* and *mbbc* quadruple mutant plants, but not Col-0 plants, showed excessive proliferation of the endogenous endophytic bacterial community

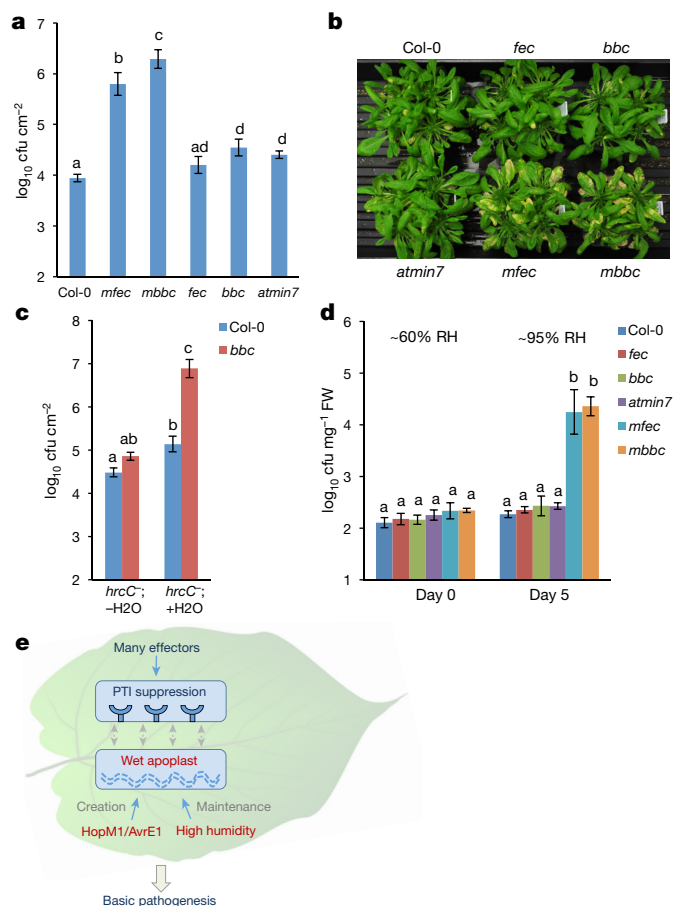


Figure 5 | Disease reconstitution experiments. **a, b**, *hrcC*[−] bacterial populations 5 days (**a**) and disease symptoms 10 days (**b**) after dip-inoculation in Col-0, *fec*, *bbc*, *atmin7*, *mfec* and *mbbc* plants grown in Redi-Earth soil. Humidity: ~95%. Statistical analysis was performed by one-way ANOVA with Tukey's test (significance set at $P \leq 0.05$). Bacterial populations indicated by different letters (a, b, c and d) are significantly different (ad is not significantly different from a or d). $n = 4$ technical replicates; data are shown as mean \pm s.d. Experiments were repeated four times. **c**, *hrcC*[−] bacterial populations in Col-0 and *bbc* leaves 3 days after infiltration with bacteria at 1×10^6 cfu ml^{−1}. The $-H_2O$ and $+H_2O$ conditions are as in Fig. 2e. Statistical analysis was performed by one-way ANOVA with Tukey's test (significance set at $P \leq 0.05$). Bacterial populations indicated by different letters (a, b and c) are significantly different (ab is not significantly different from a or b). $n = 3$ technical replicates; data are shown as mean \pm s.d. Experiments were repeated three times. **d**, Col-0, *fec*, *bbc*, *atmin7*, *mfec* and *mbbc* plants grown in *Arabidopsis* Mix soil were mock-sprayed with H_2O and kept under high humidity (~95%). On day 0 (before water spray) and day 5, total populations of the endophytic bacterial community were quantified by counting CFUs on R2A plates, after surface sterilization of leaves with 75% ethanol, leaf homogenization and serial dilutions. FW, fresh weight. Statistical analysis as in **a**. Bacterial populations indicated by different letters (a and b) are significantly different. $n = 4$ technical replicates; data are shown as mean \pm s.d. Experiments were repeated three times. **e**, A proposed new model for *Pst*-DC3000 pathogenesis in *Arabidopsis*. Dashed arrows indicate a possible interaction, at spatial and temporal scales, between 'immunosuppression' and a 'wet apoplast' during pathogenesis.

(Fig. 5d and Extended Data Table 1), in a high-humidity-dependent manner (Extended Data Fig. 8a). Furthermore, the excessive proliferation of the endophytic bacterial community was associated with mild tissue chlorosis and necrosis in some leaves (Extended Data Fig. 8b). This was noteworthy, because a recent study showed that overgrowth of a beneficial root-colonizing fungus in immunocompromised (against fungal pathogens) plants also led to harmful effects in *Arabidopsis*²⁸, illustrating a potentially common occurrence where the levels of

commensal and beneficial microbiota must be strictly controlled by the host for optimal plant health. It would be interesting in the future to conduct comprehensive *in planta* 16S rRNA amplicon-based analysis to determine whether there are also humidity-dependent changes in the composition of commensal bacterial communities in the Col-0, *mfec* and *mbbc* plants.

Discussion

Our data suggest a new conceptual framework for understanding phyllosphere–bacterial interactions (Fig. 5e). Specifically, we have identified PTI signalling and AtMIN7, presumably via vesicle trafficking, as two key components of the elusive host barrier that functions to limit excessive and potentially harmful proliferation of non-pathogenic microbes (for example, the *hrcC*[−] mutant) in the phyllosphere. Pathogenic bacteria, such as *Pst*-DC3000, have evolved T3SS effectors not only to disarm PTI signalling but also to establish an aqueous living space in a humidity-dependent manner in order to aggressively colonize the phyllosphere. The conceptual framework presented here integrates host, pathogen and environmental factors, and provides a critical insight into the basis of the strong effect of humidity on the development of numerous bacterial diseases, consistent with the 'disease triangle' dogma in plant pathology.

Prior to this study, humidity was commonly thought to promote bacterial movements on the plant surface and invasion into plant tissues. Our study, however, reveals a notable and previously unrecognized effect of high humidity on the function of bacterial effectors inside the plant apoplast. An aqueous apoplast could potentially facilitate the flow of nutrients to bacteria, promote the spread and/or egression of bacteria, and/or affect apoplastic–host–defence responses, the latter of which may underlie some of the previously observed effects of HopM1, AvrE1 and AtMIN7 on plant immunity^{21,23,25} and suggest a potential cross-talk between plant immune responses and water availability.

Most of our current knowledge on plant–pathogen and plant–microbiome interactions are derived from studies under limited laboratory conditions. This study illustrates a need for future research to consider the dynamic climate conditions in which plants and microbes live in nature in order to uncover new biological phenomena involved in host–microbe interactions. Research that unravels the molecular bases of environmental influences of disease development should help us to understand the severity, emergence and/or disappearance of infectious diseases in crop fields and natural ecosystems, especially in light of the rapidly changing drought–humidity patterns associated with global climate change.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 9 May; accepted 17 October 2016.

- Miller, S., Rowe, R. & Riedel, R. Bacterial spot, speck, and canker of tomatoes. *Ohio State University Extension Fact Sheet* HYG-3120-96 (1996).
- Pernezny, K. & Zhang, S. Bacterial speck of tomato. *University of Florida IFAS Extension PP-10* (2005).
- Schwartz, H. F. Bacterial diseases of beans. *Colorado State University Extension. Fact Sheet No. 2.913*. (2011).
- Stevens, R. B. *Plant Pathology, an Advanced Treatise*. Vol. 3, 357–429 (Academic, 1960).
- Büttner, D. & He, S. Y. Type III protein secretion in plant pathogenic bacteria. *Plant Physiol.* **150**, 1656–1664 (2009).
- Galán, J. E. & Collmer, A. Type III secretion machines: bacterial devices for protein delivery into host cells. *Science* **284**, 1322–1328 (1999).
- Asai, S. & Shirasu, K. Plant cells under siege: plant immune system versus pathogen effectors. *Curr. Opin. Plant Biol.* **28**, 1–8 (2015).
- Dou, D. & Zhou, J. M. Phytopathogen effectors subverting host immunity: different foes, similar battleground. *Cell Host Microbe* **12**, 484–495 (2012).
- Macho, A. P. & Zipfel, C. Targeting of plant pattern recognition receptor-triggered immunity by bacterial type-III secretion system effectors. *Curr. Opin. Microbiol.* **23**, 14–22 (2015).
- Asrat, S., Davis, K. M. & Isberg, R. R. Modulation of the host innate immune and inflammatory response by translocated bacterial proteins. *Cell. Microbiol.* **17**, 785–795 (2015).

11. Sperandio, B., Fischer, N. & Sansonetti, P. J. Mucosal physical and chemical innate barriers: lessons from microbial evasion strategies. *Semin. Immunol.* **27**, 111–118 (2015).
12. Gimenez-Ibanez, S., Ntoukakis, V. & Rathjen, J. P. The LysM receptor kinase CERK1 mediates bacterial perception in *Arabidopsis*. *Plant Signal. Behav.* **4**, 539–541 (2009).
13. Macho, A. P. & Zipfel, C. Plant PRRs and the activation of innate immune signaling. *Mol. Cell* **54**, 263–272 (2014).
14. Schwessinger, B. *et al.* Phosphorylation-dependent differential regulation of plant growth, cell death, and innate immunity by the regulatory receptor-like kinase BAK1. *PLoS Genet.* **7**, e1002046 (2011).
15. Tsuda, K., Sato, M., Stoddard, T., Glazebrook, J. & Katagiri, F. Network properties of robust immunity in plants. *PLoS Genet.* **5**, e1000772 (2009).
16. Yuan, J. & He, S. Y. The *Pseudomonas syringae* Hrp regulation and secretion system controls the production and secretion of multiple extracellular proteins. *J. Bacteriol.* **178**, 6399–6402 (1996).
17. Cunnac, S. *et al.* Genetic disassembly and combinatorial reassembly identify a minimal functional repertoire of type III effectors in *Pseudomonas syringae*. *Proc. Natl Acad. Sci. USA* **108**, 2975–2980 (2011).
18. Hirano, S. S. & Upper, C. D. Population biology and epidemiology of *Pseudomonas syringae*. *Annu. Rev. Phytopathol.* **28**, 155–177 (1990).
19. Fan, J., Crooks, C. & Lamb, C. High-throughput quantitative luminescence assay of the growth in planta of *Pseudomonas syringae* chromosomally tagged with *Photobacterium luminescens* luxCDABE. *Plant J.* **53**, 393–399 (2008).
20. Badel, J. L., Shimizu, R., Oh, H. S. & Collmer, A. A *Pseudomonas syringae* pv. *tomato* avrE1/hopM1 mutant is severely reduced in growth and lesion formation in tomato. *Mol. Plant Microbe Interact.* **19**, 99–111 (2006).
21. DebRoy, S., Thilmony, R., Kwack, Y. B., Nomura, K. & He, S. Y. A family of conserved bacterial effectors inhibits salicylic acid-mediated basal immunity and promotes disease necrosis in plants. *Proc. Natl Acad. Sci. USA* **101**, 9927–9932 (2004).
22. Nomura, K. *et al.* A bacterial virulence protein suppresses host innate immunity to cause plant disease. *Science* **313**, 220–223 (2006).
23. Xin, X. F. *et al.* *Pseudomonas syringae* effector Avrulence protein E localizes to the host plasma membrane and down-regulates the expression of the *NONRACE-SPECIFIC DISEASE RESISTANCE1/HARPIN-INDUCED1-LIKE13* gene required for antibacterial immunity in *Arabidopsis*. *Plant Physiol.* **169**, 793–802 (2015).
24. Wei, C. F. *et al.* A *Pseudomonas syringae* pv. *tomato* DC3000 mutant lacking the type III effector HopQ1-1 is able to cause disease in the model plant *Nicotiana benthamiana*. *Plant J.* **51**, 32–46 (2007).
25. Nomura, K. *et al.* Effector-triggered immunity blocks pathogen degradation of an immunity-associated vesicle traffic regulator in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **108**, 10774–10779 (2011).
26. Tanaka, H., Kitakura, S., De Rycke, R., De Groodt, R. & Friml, J. Fluorescence imaging-based screen identifies ARF GEF component of early endosomal trafficking. *Curr. Biol.* **19**, 391–397 (2009).
27. Whalen, M., C., Innes, R. W., Bent, A. F. & Staskawicz, B. J. Identification of *Pseudomonas syringae* pathogens of *Arabidopsis* and a bacterial locus determining avirulence on both *Arabidopsis* and soybean. *Plant Cell* **3**, 49–59 (1991).
28. Hiruma, K. *et al.* Root endophyte *Colletotrichum tofieldiae* confers plant fitness benefits that are phosphate status dependent. *Cell* **165**, 464–474 (2016).
29. Hauck, P., Thilmony, R. & He, S. Y. A *Pseudomonas syringae* type III effector suppresses cell wall-based extracellular defense in susceptible *Arabidopsis* plants. *Proc. Natl Acad. Sci. USA* **100**, 8577–8582 (2003).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank He laboratory members for insightful discussions and constructive suggestions. We thank J. Kremer for help with setting up real-time disease imaging experiments and advice on 16S rRNA amplicon sequencing, K. Sugimoto for providing tomato plants (cv. Castlemart), and C. Thireault for technical help. This project was supported by funding from Gordon and Betty Moore Foundation (GBMF3037), National Institutes of Health (GM109928) and the Department of Energy (the Chemical Sciences, Geosciences, and Biosciences Division, Office of Basic Energy Sciences, Office of Science; DE-FG02-91ER20021 for infrastructural support). C.Z. acknowledges support from The Gatsby Charitable Foundation.

Author Contributions X.-F.X., K.N. and S.Y.H. designed the experiments. K.A. performed the *Pst*-DC3000-*lux* imaging experiment. A.C.V. performed biological repeats of bacterial infection experiments shown in Fig. 1a. J.Y. characterized an unpublished plant mutant line. X.-F.X. and K.N. performed all other experiments, including bacterial infections, protein blotting and generation of *Arabidopsis* *mfec* and *mbbc* mutant lines. F.B. and C.Z. contributed unpublished plant mutant materials. J.H.C. contributed unpublished *Pst*-DC3000 effector constructs. X.-F.X. and S.Y.H. wrote the manuscript with input from all co-authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.Y.H. (hes@msu.edu).

Reviewer Information Nature thanks G. Beattie, S. Lindow, J.-M. Zhou and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

Plant materials and bacterial strains. *Arabidopsis thaliana* plants were grown in *Arabidopsis* Mix soil (equal parts of SUREMIX (Michigan Grower Products Inc.), medium vermiculate and perlite; autoclaved once) or Redi-Earth soil (Sun Gro Horticulture) in environmentally controlled growth chambers, with relative humidity at 60%, temperature at 22 °C and a 12 h light–12 h dark cycle. Five-week-old plants were used for bacterial inoculation and disease assays.

The *bak1-5 bkk1-1 cerk1* mutant plant was generated by crossing the *bak1-5 bkk1-1* mutant¹⁴ with the *cerk1* mutant³⁰. PCR-based genotyping was performed in F₂ progeny to obtain a homozygous triple mutant. The *npr1-6* (Fig. 4a) mutant was the SAIL_708_F09 line ordered from the *Arabidopsis* Biological Resource Center, and confirmed to be a knockout mutant and defective in salicylic acid signalling (Extended Data Fig. 6).

Bacterial disease assays. Syringe-infiltration and dip-inoculation were performed in this study. Briefly, *Pst*-DC3000 and mutant strains were cultured in modified Luria–Bertani medium (10 g l⁻¹ tryptone, 6 g l⁻¹ yeast extract, 1.5 g l⁻¹ KH₂PO₄, 0.6 g l⁻¹ NaCl, and 0.4 g l⁻¹ MgSO₄·7H₂O) containing 100 mg l⁻¹ rifampicin (and/or other antibiotics if necessary) at 28 °C to an OD₆₀₀ of 0.8–1.0. Bacteria were collected by centrifugation and re-suspended in sterile water. Cell density was adjusted to OD₆₀₀ = 0.2 (approximately 1 × 10⁸ cfu ml⁻¹). For syringe-infiltration, bacterial suspension was further diluted to cell densities of 1 × 10⁵–1 × 10⁶ cfu ml⁻¹. Unless stated otherwise, infiltrated plants were first kept under ambient humidity for 1–2 h for water to evaporate, and after the plant leaves returned to pre-infiltration appearance, plants were kept under high humidity (approximately 95%; by covering plants with a clear plastic dome) or other specified humidity settings for disease to develop. For dip-inoculation, plants were dipped in the bacterial suspension of OD₆₀₀ = 0.2, with 0.025% Silwet L-77 added, and then kept under high humidity (approximately 95%) immediately for disease to develop.

Different humidity settings were achieved by placing a clear plastic dome over a flat (in which plants are grown) with different degrees of opening. A humidity–temperature Data Logger (Lascar) was placed inside the flat to record the humidity and/or temperature over the period of disease assay.

For quantification of *Pst*-DC3000 bacterial populations, *Arabidopsis* leaves were surface-sterilized in 75% ethanol and rinsed in sterile water twice. Leaf disks were taken using a cork borer (9.5 mm in diameter) and ground in sterile water. Colony-forming units were determined by serial dilutions and plating on Luria–Marine plates containing 100 mg l⁻¹ rifampicin. One technical replicate consists of one or two leaf disks and 3 or 4 technical replicates (that is, from at least 3 or 4 leaves) were included in each biological experiment. Experiments were repeated at least three times.

CRISPR–Cas9-mediated mutation of the *AtMIN7* gene. The one-plasmid CRISPR–Cas9 cloning system³¹ was used to mutate *AtMIN7* into the *fls2 efr cerk1* and *bak1-5 bkk1-1 cerk1* plants. *AtMIN7*-single guide (sg)RNA primers containing target mutation regions were as follows, with *AtMIN7* sequence underlined.

AtMIN7-sgRNA-F, GATTGATCATTTGGAAGGGGATCC; *AtMIN7*-sgRNA-R, AAACGGATCCCCTTCCAAATGATC.

The constructs containing *AtMIN7*-sgRNA and *Cas9* were cloned in pCAM-BIA1300, which were then mobilized into *Agrobacterium tumefaciens* for plant transformation. For genotyping of *AtMIN7*-mutated lines, total DNA was extracted from individual lines and the regions containing the CRISPR target sites were amplified by PCR using the following primers:

AtMIN7-sgRNA-F2, GATGCTGCTTTGGATTGTCTTC; *AtMIN7*-sgRNA-R2, AATGGCTCCCCATGCACTGCGATA.

For genotyping, the PCR products were digested with the *Bam*HI restriction enzyme and plant lines showing an (partially or completely) uncut band were chosen. The PCR products of putative homozygous T₂ lines, identified based on a lack of cutting by *Bam*HI, were sequenced. The lines showing a frame-shift mutation and an absence of the *Cas9* gene based on PCR using the following primers were identified as homozygous lines. The T₃ and T₄ progeny of homozygous lines were used for disease assays. Primers for PCR-amplifying *Cas9* gene were as follows: *Cas9*-F, CCAGCAAGAAATTCAAGGTGC; *Cas9*-R, GCACCAGCTGGATGAACAGCTT.

Imaging of bacterial colonization with luciferase assay. Four-week-old *Arabidopsis* Col-0 plants were dip-inoculated with the *Pst*-DC3000 or *Pst*-DC3000–*lux* strain¹⁹. The infected plants were fully covered with a plastic dome to maintain high humidity. Leaves were excised from the infected plants 2 days after inoculation and the light signals were captured by a charge-coupled device (CCD) using ChemiDoc™ MP system (Bio-Rad).

AtMIN7 protein blot. *Arabidopsis* leaves were syringe-infiltrated with bacteria or H₂O and kept under high humidity (approximately 95%) for 24 h. Leaf disks were homogenized in 2 × SDS buffer, boiled for 5 min and centrifuged at 10,000g for 1 min. Supernatants containing the total protein extracts were subjected to separation by SDS–polyacrylamide gel electrophoresis (PAGE). An AtMIN7 antibody²² was used for western blotting to detect the AtMIN7 protein. Uncropped blot and gel images are included in Supplementary Fig. 1.

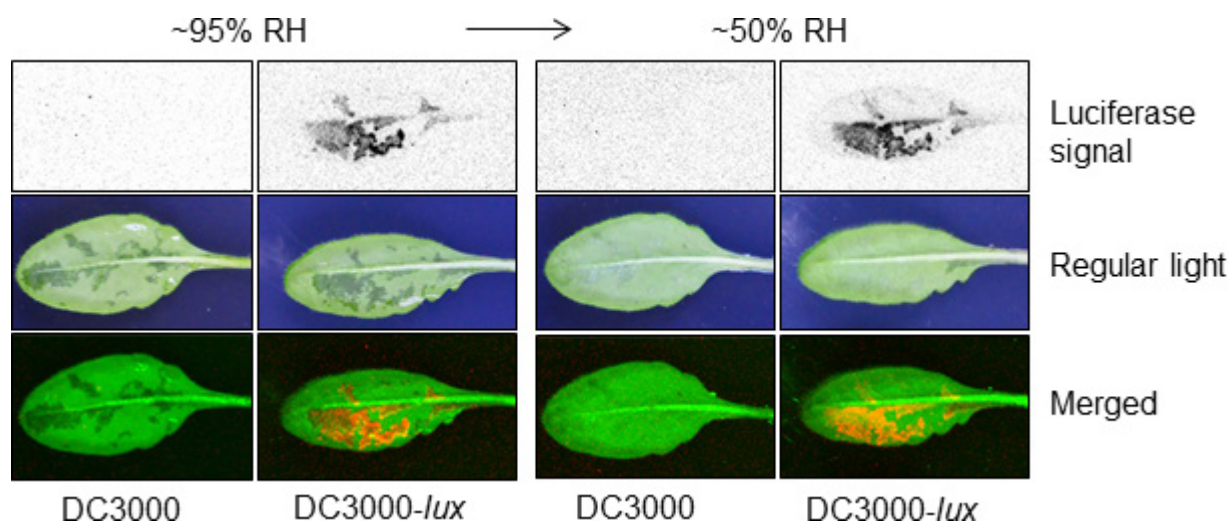
Bacterial community quantification. Five-week-old plants were sprayed with H₂O and covered with a plastic dome to keep high humidity (approximately 95%) for 5 days. To quantify the endophytic bacterial community, leaves were detached, sterilized in 75% ethanol for 1 min (Extended Data Fig. 9) and rinsed in sterile water twice. Leaves were weighed and ground in sterile water using a TissueLyser (Qiagen; at the frequency of 30 times per second for 1 min) in the presence of 3-mm Zirconium-oxide grinding beads (Glen Mills; 5 beads in each tube). After serial dilutions, bacterial suspensions were plated on R2A plates (Teknova), which were kept at 22 °C for 4 days before colonies were counted. Colony-forming units were normalized to tissue fresh weight.

16S rRNA amplicon sequence analysis of the endophytic bacterial community. The Col-0, *mfec* and *mbbc* plants were sprayed with water and kept under high humidity (approximately 95%) for 5 days. Leaves were surface-sterilized in 75% ethanol for 1 min and rinsed in sterile water twice. Leaves from four plants were randomly selected (2 leaves from each plant; 8 leaves in total) and were divided into 4 tubes (2 leaves in each tube) and ground in sterile water. Bacterial suspensions were diluted (Col-0 samples were diluted to 10⁻³ and *mfec* and *mbbc* samples were diluted to 10⁻⁵) and, for each genotype, 15 µl suspension from each tube of the right dilution (10⁻³ dilution for Col-0 and 10⁻⁵ for *mfec* and *mbbc*) were pooled together and plated on R2A plates, which were kept at 22 °C for 4 days. Fifty colonies from each genotype were randomly picked, genomic DNA was extracted and PCR was performed with AccuPrime high-fidelity Taq DNA polymerase (Invitrogen) and primers 799F/1392R (ref. 32) to amplify bacterial 16S rRNA gene. The PCR product was sequenced and taxonomy of each bacterium (family level) was determined by Ribosomal Database Project at Michigan State University (<https://rdp.cme.msu.edu/>)³³.

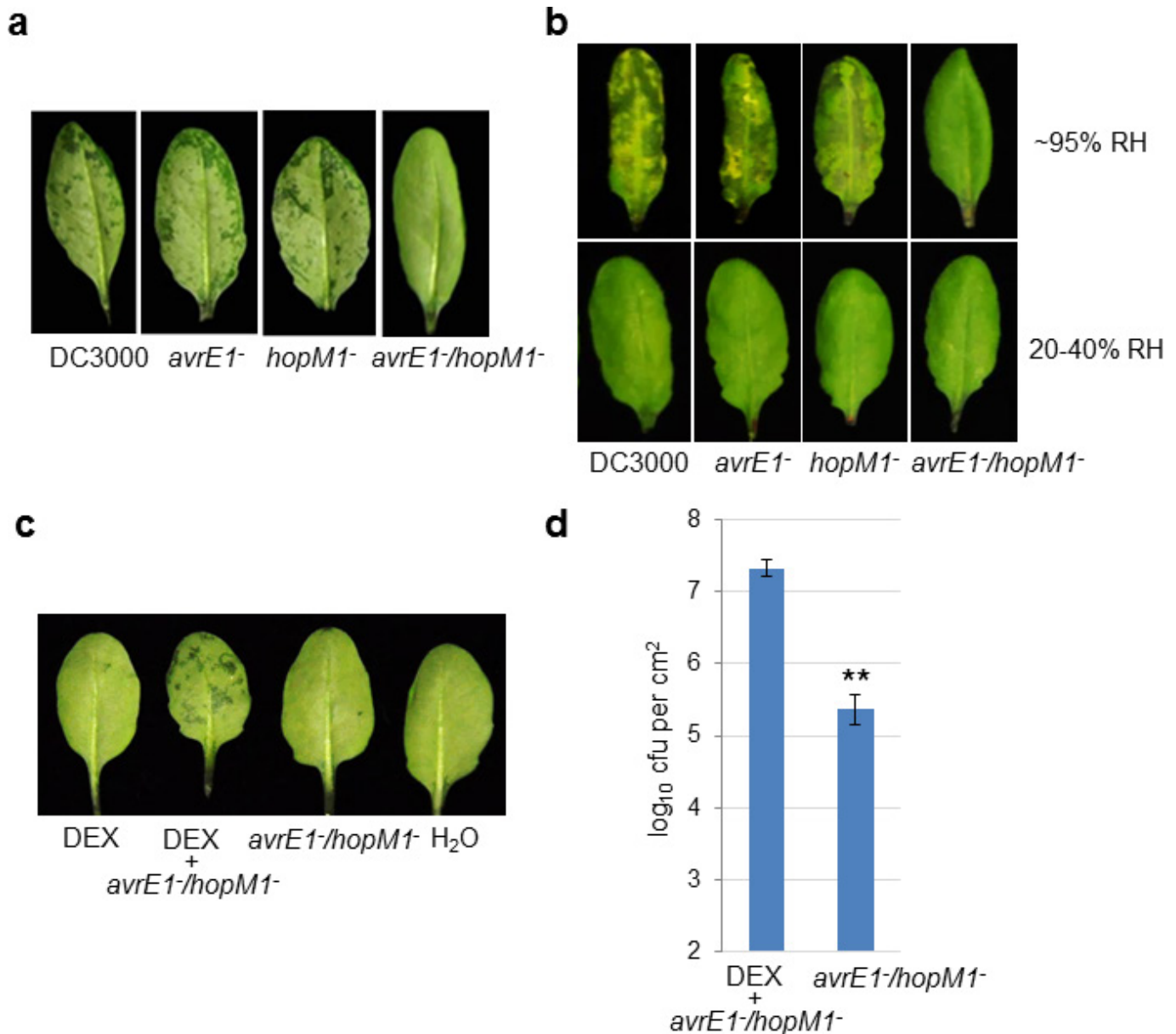
Data analysis, statistics and experimental repeats. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment. The specific statistical method used, the sample size and the results of statistical analyses are described in the relevant figure legends. Sample size was determined based on experimental trials and with consideration of previous publications on similar experiments to allow for confident statistical analyses. The Student's two-tailed *t*-test was performed for comparison of means between two data points. A one-way or two-way ANOVA with Tukey's test was used for multiple comparisons within a data set, with significance set to a *P* value ≤ 0.05. ANOVA analysis was performed with GraphPad Prism software.

Data Availability. The bacterial 16S rRNA sequences in Extended Data Table 1 have been deposited in the National Center for Biotechnology Information (NCBI) GenBank database under accession numbers KX959313–KX959462. Other data that support the findings of this study are available from the corresponding author upon request.

30. Miya, A. *et al.* CERK1, a LysM receptor kinase, is essential for chitin elicitor signaling in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **104**, 19613–19618 (2007).
31. Feng, Z. *et al.* Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **111**, 4632–4637 (2014).
32. Bai, Y. *et al.* Functional overlap of the *Arabidopsis* leaf and root microbiota. *Nature* **528**, 364–369 (2015).
33. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73**, 5261–5267 (2007).

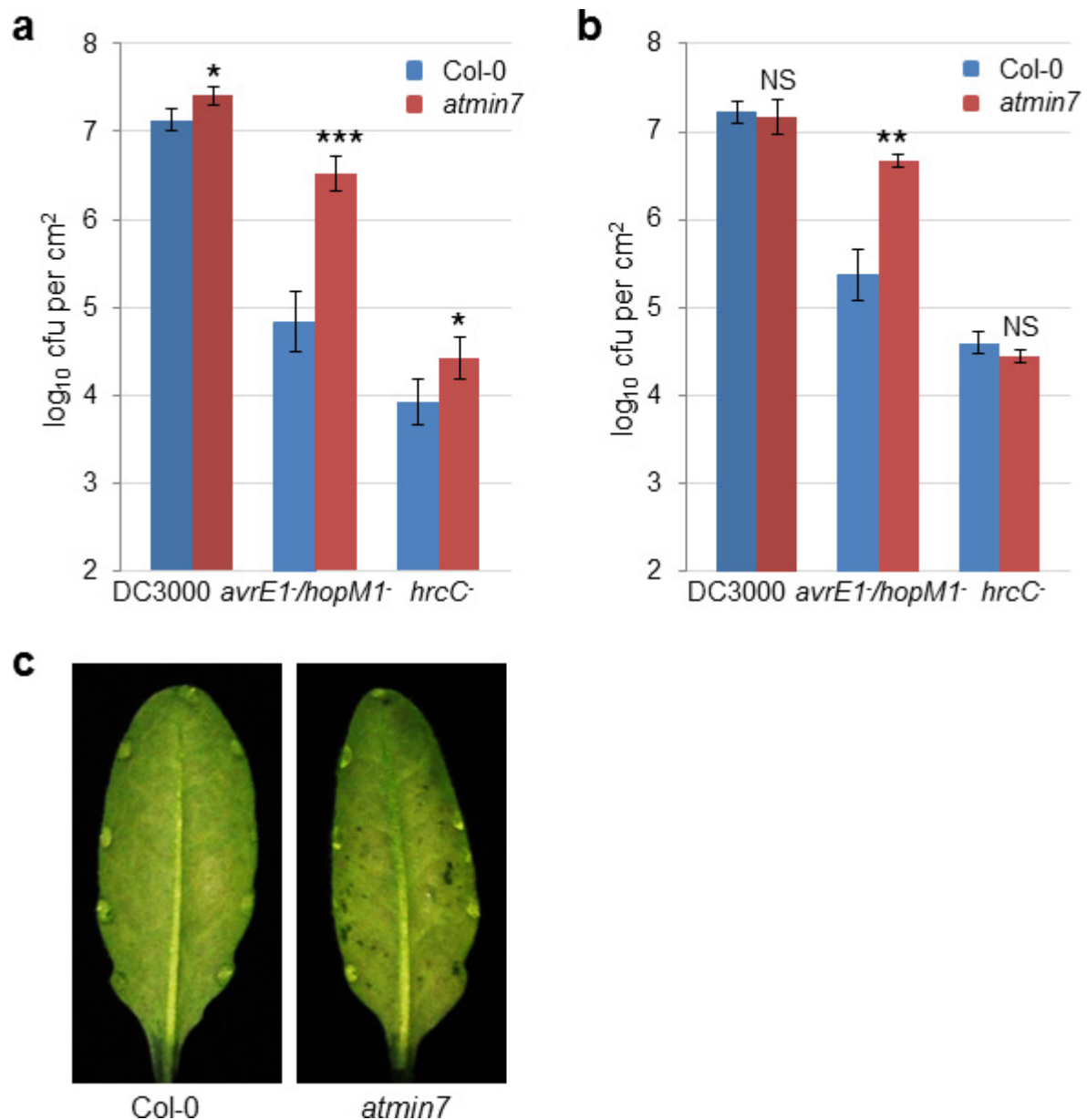


Extended Data Figure 1 | Water soaking does not affect luminescence signal. Col-0 plants were dip-inoculated with bacteria at 2×10^8 cfu ml⁻¹ and kept under high humidity (approximately 95%) for 2 days. Imaging was performed in the same way as in Fig. 1g. Water-soaked leaves were air-dried for about 2 h and imaged again (right panel). Images are representative of leaves from more than four plants.



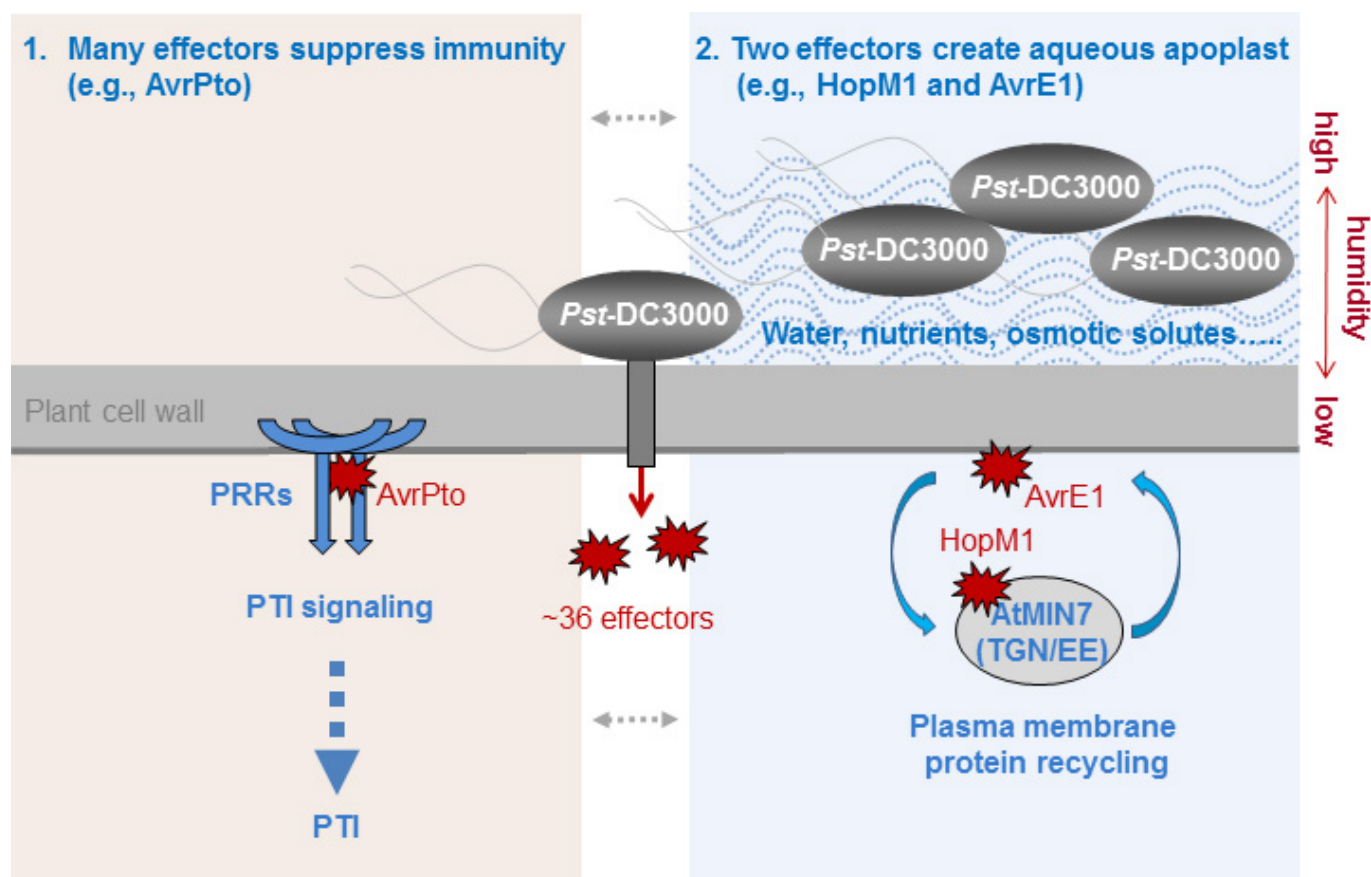
Extended Data Figure 2 | Humidity dependence of HopM1/AvrE1 function and restoration of the virulence of the *avrE1*⁻/*hopM1*⁻ mutant in 6×His:HopM1 transgenic plants. **a, b,** The virulence of the *avrE1*⁻/*hopM1*⁻ mutant is insensitive to humidity settings. **a,** Col-0 plants were syringe-infiltrated with indicated bacteria at 2×10^5 cfu ml⁻¹. Inoculated plants were kept under high (approximately 95%) humidity, and images were taken 24 h after infiltration. **b,** Col-0 plants were syringe-infiltrated with *Pst*-DC3000, the *avrE1*⁻ mutant, the *hopM1*⁻ mutant or the *avrE1*⁻/*hopM1*⁻ mutant at 2×10^5 cfu ml⁻¹. Inoculated plants were kept under high (approximately 95%) or low (20–40%) humidity. Images

were taken 3 days after inoculation. Images are representative of leaves from more than four plants. **c, d,** The 6×His:HopM1 transgenic plants²² were infiltrated with 0.1 nM dexamesathone (DEX), the *avrE1*⁻/*hopM1*⁻ mutant (at 1×10^5 cfu ml⁻¹) or both. H₂O was infiltrated as control. Infiltrated plants were kept at high humidity (approximately 95%). Leaf images were taken 24 h after infiltration (**c**) and bacterial populations were determined 3 days after infiltration (**d**). Significant difference was determined by Student's *t*-test; (two-tailed); *** $p = 1.03 \times 10^{-5}$, $n = 6$ technical replicates from three independent experiments ($n = 2$ in each experiment); data are shown as mean \pm s.d.



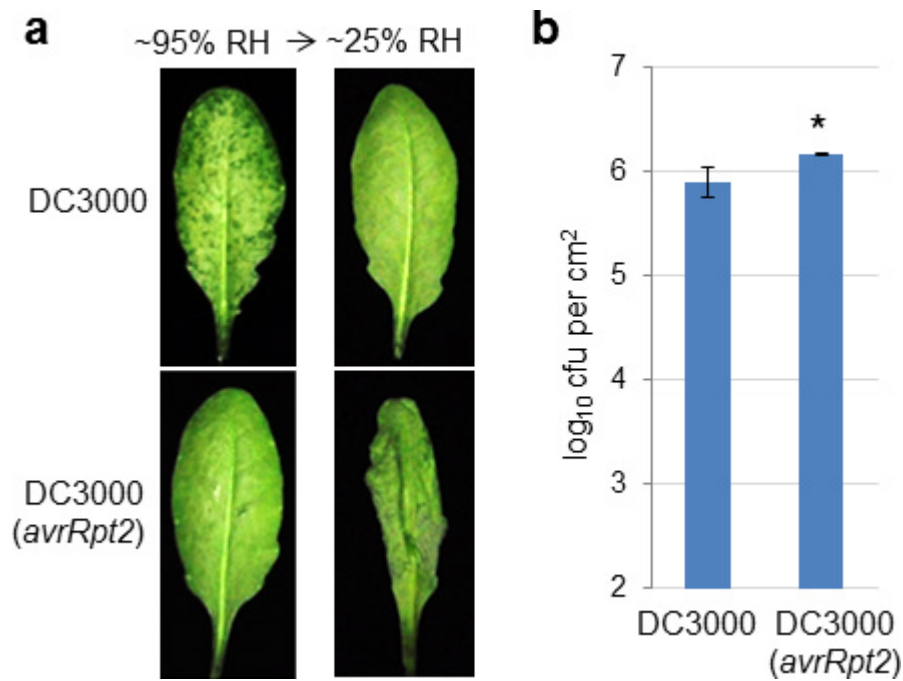
Extended Data Figure 3 | Bacterial multiplication and water soaking in Col-0 and the *atmin7* mutant. **a**, The Col-0 and *atmin7* plants were dip-inoculated with *Pst*-DC3000, the *avrE1/hopM1*⁻ mutant or the *hrcC*⁻ mutant at 1×10^8 cfu ml⁻¹. Bacterial populations were determined 4 days after inoculation. Significant difference between Col-0 and *atmin7* plants was determined by Student's *t*-test (two-tailed); * $P = 1.61 \times 10^{-2}$ and 3.12×10^{-2} for DC3000 and *hrcC*⁻, respectively; *** $P = 1.41 \times 10^{-4}$ for *avrE1/hopM1*⁻. $n = 4$ technical replicates; data are shown as mean \pm s.d. Experiments were repeated three times. **b**, **c**, The Col-0 and *atmin7* plants were syringe-infiltrated with *Pst*-DC3000, the *avrE1/hopM1*⁻ mutant

or the *hrcC*⁻ mutant at 1×10^6 cfu ml⁻¹. Bacterial populations were determined 3 days after inoculation (**b**) and leaf images were taken 38 h after infiltration with the *avrE1/hopM1*⁻ mutant strain to show water soaking in *atmin7* leaves (**c**). Significant difference between Col-0 and *atmin7* plants was determined by Student's *t*-test (two-tailed); ** $P = 1.63 \times 10^{-3}$ for *avrE1/hopM1*⁻; NS, not significant ($P = 0.72$ and 0.14 for DC3000 and *hrcC*⁻, respectively). $n = 3$ technical replicates; data are shown as mean \pm s.d. Experiments were repeated three times. Images are representative of leaves from more than four plants.



Extended Data Figure 4 | A working model showing the function of HopM1 and AvrE1 in creating aqueous apoplasts. *Pst*-DC3000 delivers a total of 36 effectors into the plant cell. Many effectors, including AvrPto, appear to suppress pattern-triggered immunity (PTI). AvrPto inhibits pattern recognition receptor (PRR) function⁸. Two conserved effectors,

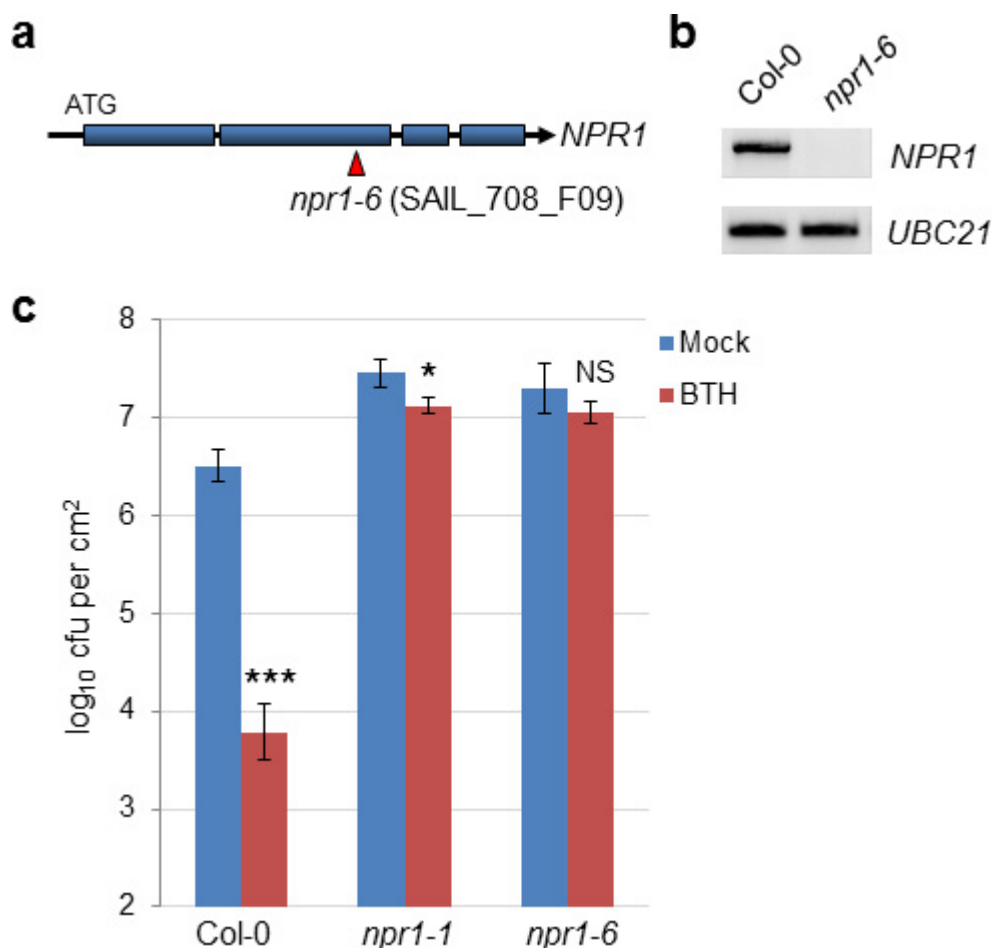
HopM1 and AvrE1, create an aqueous apoplast in the presence of bacterial infection in a humidity-dependent manner. AvrE1 is localized to the host plasma membrane²³. HopM1 targets AtMIN7 (an ARF-GEF protein) in the TGN/EE, which is involved in recycling of plasma membrane proteins²⁶.



Extended Data Figure 5 | Water-soaking is blocked during ETI.

a, Col-0 leaves were syringe-infiltrated with *Pst*-DC3000 (1×10^6 cfu ml⁻¹) or *Pst*-DC3000(*avrRpt2*) (1×10^7 cfu ml⁻¹). Plants were kept under high humidity (approximately 95%) for 24 h to observe water soaking and then shifted to low humidity (approximately 25%) for 2 h to observe ETI-associated tissue collapse. Images were taken before and after low humidity

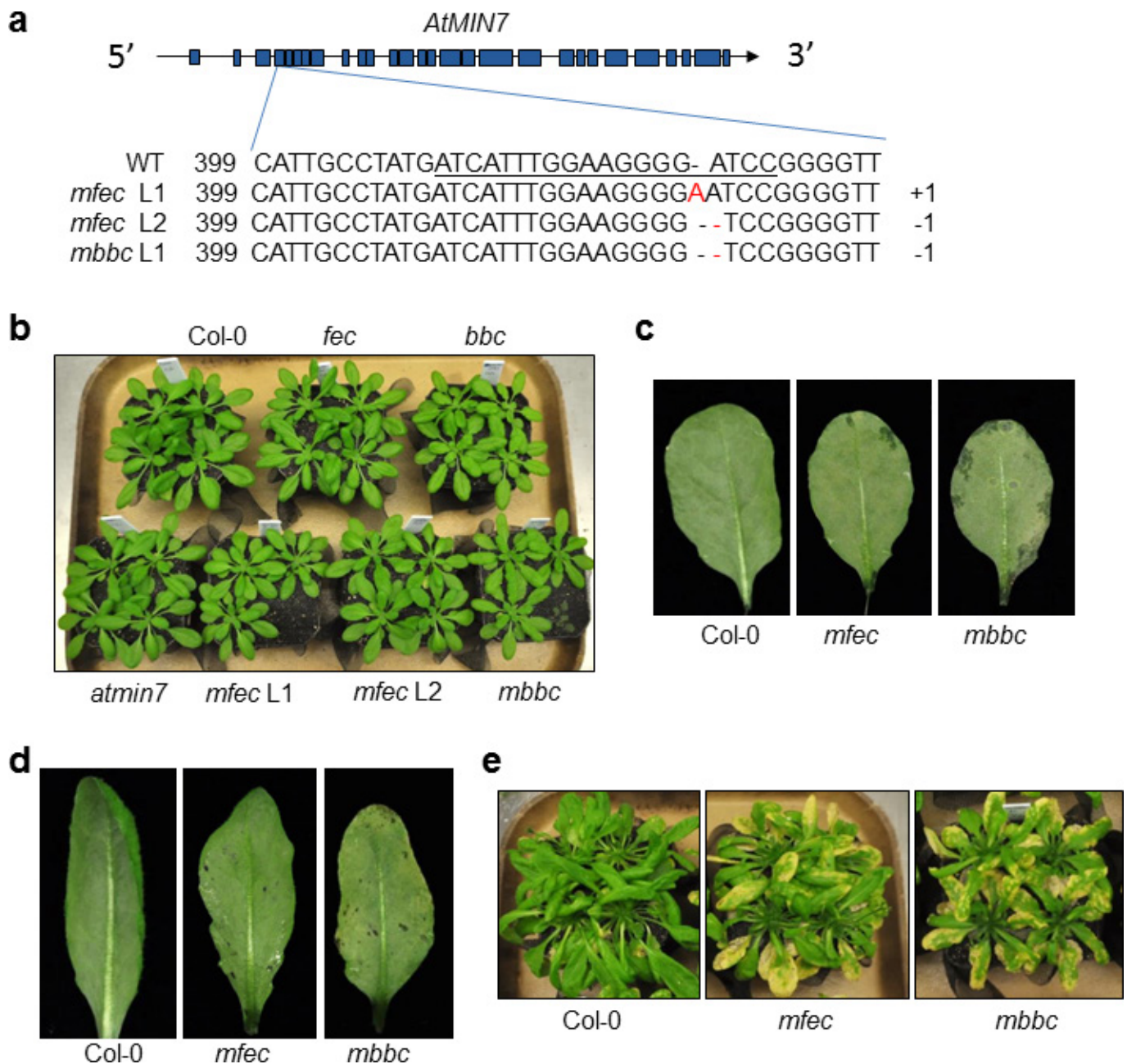
exposure (**a**) and bacterial populations were determined 24 h after infiltration to show similar population levels (**b**). Significant difference in bacterial population was determined by Student's *t*-test (two-tailed); **P* = 0.033. *n* = 3 technical replicates; data are shown as mean ± s.d. Experiments were repeated three times. This is an experimental replicate of Fig. 3b, c (without *rps2* mutant plants).



Extended Data Figure 6 | Characterization of the *npr1-6* mutant.

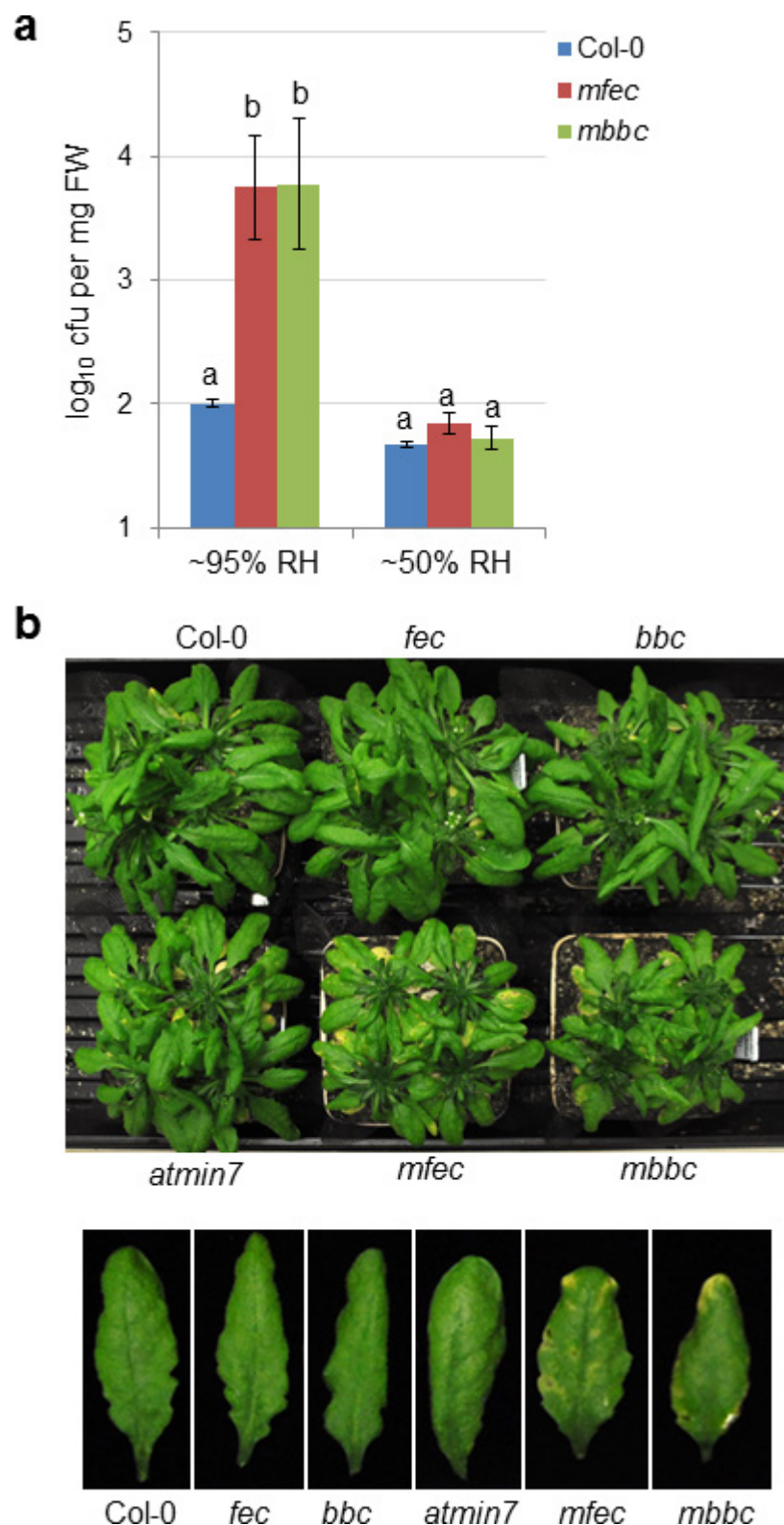
a, A diagram showing the T-DNA insertion site in the *npr1-6* mutant. Blue boxes indicate exons in the *NPR1* gene. **b**, RT-PCR results showing that the *npr1-6* line cannot produce the full-length *NPR1* transcript. Primers used (*NPR1* sequence is underlined): *NPR1*-F: agaattcATGGACACCACCATTGATGGA; *NPR1*-R: agtcgacCCGACGACGATGAGAGARTTTAC; *UBC21*-F: TCAAATG GACCGCTCTTATC; *UBC21*-R: TCAAATGGACCGCTCTTATC. Uncropped gel images are shown in Supplementary Fig. 1. **c**, The *npr1-6* line, similar to

npr1-1, is greatly compromised in benzothiadiazole (BTH)-mediated resistance to *Pst*-DC3000 infection. The Col-0, *npr1-1* and *npr1-6* plants were sprayed with 100 μ M BTH and dip-inoculated with *Pst*-DC3000 at 1×10^8 cfu ml⁻¹ 24 h later. Bacterial populations were determined 3 days after inoculation. Significant difference between mock and BTH treatment was determined by Student's *t*-test (two-tailed); * $P = 0.027$; *** $P = 1.6 \times 10^{-4}$; NS, not significant ($P = 0.19$). $n = 3$ technical replicates; data are shown as mean \pm s.d. Experiments were repeated three times.



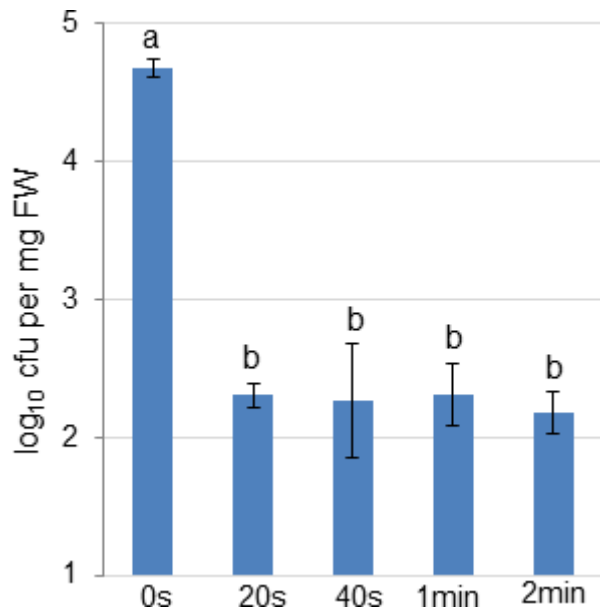
Extended Data Figure 7 | Construction and characterization of the *mfec* and *mbbc* quadruple mutants. **a**, CRISPR-Cas9-mediated mutations in the 4th exon of the *AtMIN7* gene (exons indicated by blue boxes) in the quadruple mutant lines used in this study. The underlined sequence in the wild type (WT) indicates the region targeted by sgRNA. The number 399 indicates the nucleotide position in the *AtMIN7* coding sequence. +1 and -1 indicate frame shifts in the mutant lines. **b**, Col-0 and various mutants used in this study have similar growth, development and morphology. Four-week-old plants are shown. **c**, The *mfec* and *mbbc* plants show a tendency to develop sporadic water soaking under high humidity. Five-week-old regularly-grown (around 60% relative humidity) Col-0, *mfec* and *mbbc* plants were shifted to high humidity (approximately 95%)

overnight and images of mature leaves were taken after the high humidity incubation. **d**, Even leaves of *mfec* and *mbbc* plants that do not have sporadic water soaking have a tendency to develop some water soaking after *hrcC*⁻ inoculation. Five-week-old Col-0, *mfec* and *mbbc* plants were dip-inoculated with *hrcC*⁻ at 1×10^8 cfu ml⁻¹, and kept under high humidity (approximately 95%). Leaf images were taken 2 days after inoculation. Images are representative of leaves from at least four plants. **e**, The non-pathogenic *hrcC*⁻ mutant causes significant necrosis and chlorosis in the quadruple mutant plants. Col-0, *mfec* and *mbbc* plants were dip-inoculated with the *hrcC*⁻ strain at 1×10^8 cfu ml⁻¹. Images were taken 9 days after inoculation. This is one of the four independent experimental replications of the results presented in Fig. 5b.



Extended Data Figure 8 | Multiplication of endophytic phyllosphere bacterial community. **a**, An increase in the endophytic bacterial community in *mfec* and *mbbc* plants depends on high humidity. Col-0, *mfec* and *mbbc* plants were either sprayed with H₂O and kept under high humidity (approximately 95%) or low humidity (around 50%). On day 5, total populations of the endophytic bacterial community were quantified. Statistical analysis was performed by one-way ANOVA with Tukey's test (significance set at $P \leq 0.05$). Bacterial populations indicated by different

letters (a and b) are significantly different. $n = 4$ technical replicates; data are shown as mean \pm s.d. Experiments were repeated three times. **b**, Mild chlorosis and necrosis in leaves is associated with increased endophytic bacterial community level in the *mfec* and *mbbc* quadruple mutant plants. Plants were sprayed with H₂O and kept under high (approximately 95%) humidity. Images were taken 10 days after spraying. Individual leaves are enlarged and shown in the lower panel, showing mild chlorosis and necrosis in some of the *mfec* and *mbbc* leaves.



Extended Data Figure 9 | Validation of 1 min as an effective surface sterilization time. Five-week-old Col-0 plants were sprayed with H₂O and kept under high humidity (approximately 95%) for 5 days. Leaves were detached, surface sterilized in 75% ethanol for 20 s, 40 s, 1 min or 2 min and then rinsed in sterile water twice. No sterilization (0 s) was used as control. Leaves were ground in sterile water and bacterial numbers were determined by serial dilutions and counting of colony-forming units on R2A plates. Statistical analysis was performed by one-way ANOVA with a Tukey's test (significance set at $P \leq 0.05$). Bacterial populations indicated by different letters (that is, a and b) are significantly different. $n = 4$ technical replicates; data are shown as mean \pm s.d. Experiments were repeated twice with similar results.

Extended Data Table 1 | Endophytic bacterial taxa in Col-0, *mfec* and *mbbc* plants

Order/Family	Col-0	<i>mfec</i>	<i>mbbc</i>
Bacillales			
Paenibacillaceae	15 (30%)	ND	ND
Burkholderiales			
Comamonadaceae	8 (16%)	12 (24%)	9 (18%)
Burkholderiaceae	4 (8%)	1 (2%)	22 (44%)
Alcaligenaceae	3 (6%)	19 (38%)	12 (24%)
Flavobacteriales			
Flavobacteriaceae	6 (12%)	1 (2%)	1 (2%)
Xanthomonadales			
Xanthomonadaceae	4 (8%)	9 (18%)	ND
Sphingomonadales			
Sphingomonadaceae	3 (6%)	ND	1 (2%)
Sphingobacteriales			
Sphingobacteriaceae	3 (6%)	ND	ND
Chitinophagaceae	1 (2%)	ND	ND
Rhizobiales			
Rhizobiaceae	2 (4%)	5 (10%)	ND
Cytophagales			
Cytophagaceae	1 (2%)	ND	ND
Pseudomonadales			
Pseudomonadaceae	ND	1 (2%)	5 (10%)
Actinomycetales			
Microbacteriaceae	ND	2 (4%)	ND

ND, not detected. See Methods for 16S rRNA amplicon sequencing procedures.

The mechanism of force transmission at bacterial focal adhesion complexes

Laura M. Faure^{1*}, Jean-Bernard Fiche^{2*}, Leon Espinosa^{1*}, Adrien Ducret^{1,3}, Vivek Anantharaman⁴, Jennifer Luciano¹, Sébastien Lhospice¹, Salim T. Islam¹, Julie Tréguier¹, Mélanie Sotes¹, Erkin Kuru⁵, Michael S. Van Nieuwenhze⁶, Yves V. Brun³, Olivier Théodoly⁷, L. Aravind⁴, Marcelo Nollmann² & Tâm Mignot¹

Various rod-shaped bacteria mysteriously glide on surfaces in the absence of appendages such as flagella or pili. In the *deltaproteobacterium* *Myxococcus xanthus*, a putative gliding motility machinery (the Agl–Glt complex) localizes to so-called focal adhesion sites (FASs) that form stationary contact points with the underlying surface. Here we show that the Agl–Glt machinery contains an inner-membrane motor complex that moves intracellularly along a right-handed helical path; when the machinery becomes stationary at FASs, the motor complex powers a left-handed rotation of the cell around its long axis. At FASs, force transmission requires cyclic interactions between the molecular motor and the adhesion proteins of the outer membrane via a periplasmic interaction platform, which presumably involves contractile activity of motor components and possible interactions with peptidoglycan. Our results provide a molecular model of bacterial gliding motility.

Certain rod-shape bacteria move along their long axes in the absence of extracellular appendages such as flagella or pili, in a process called gliding motility¹. In *Myxococcus xanthus*, gliding is mediated by bacterial FASs. During gliding motility, FASs assemble at the leading cell pole and retain a fixed position relative to the surface until they disassemble at the lagging cell pole¹. FASs contain a molecular machinery, the Agl–Glt (adventurous gliding and gliding transducer) complex, which comprises more than fourteen proteins^{2–4}. Genetic analysis has suggested that this machinery is formed by two membrane-associated systems: (i) a putative three-protein TolQR-like proton motive force (PMF)-driven channel comprising AglR, AglQ and AglS (the suspected energy-producing system)^{3,5,6}; and (ii) a putative eleven-protein integral envelope-associated complex (containing GltA–K) that interacts with the Agl system^{2–4}. These machineries are further connected to a cytosolic protein complex formed by the AglZ protein, the Ras-like G-protein MglA and the MreB actin cytoskeleton (hereafter called the MreB complex⁷; Fig. 1a). This MreB complex recruits, and promotes the assembly of the Agl–Glt complex (Fig. 1a). Directional movements of the Agl–Glt complex from the leading towards the lagging cell pole have been suggested to propel the cell forward^{3,8}, but how these movements may be transduced into cell movement remains unknown^{5,9}. Here, we reveal the functional architecture of the Agl–Glt complex and establish how its activity is transduced to the contact surface across the highly structured layers of the cell envelope.

Intracellular helically moving engines drive propulsion

We quantitatively characterized the dynamic behaviour of FASs by analysing the motions of AglZ–YFP (yellow fluorescent protein)-containing complexes using total internal reflection fluorescence microscopy (TIRFM). Cells attached to a chitosan-coated surface alternated between motile and non-motile states. In these conditions, we were able to capture the movement of AglZ–YFP clusters over extended

periods of time with high temporal resolution. Two main AglZ–YFP cluster populations were observed: static and dynamic (Fig. 1b; blue and orange, respectively). Motile cells on chitosan exhibited at least one static AglZ–YFP cluster, indicating that a single static cluster is necessary and sufficient for cell propulsion. We also observed dynamic AglZ–YFP clusters. These clusters tended to form at one cell pole and migrate directionally towards the opposite pole (Fig. 1b and Extended Data Fig. 1a). On average, clusters formed every minute and moved at a constant velocity ($3.2 \pm 0.9 \mu\text{m min}^{-1}$, $n = 227$) over distances of $1.5 \pm 1 \mu\text{m}$ ($n = 203$), often becoming dissociated upon reaching the opposite pole. Cluster speeds varied within and between cells (Extended Data Fig. 1b), possibly because of variations in the number of motor units in a cluster (see below) and changes in PMF levels between cells^{3,6}. Dynamic clusters are likely to represent unattached motility complexes because: (i) in motile cells, they were detected only if a fixed cluster was also present; and (ii) in most cells the transition from a non-motile to a motile state coincided with cluster immobilization ($>85\%$, $n = 34$; Fig. 1b, orange–blue cluster). To characterize this behaviour quantitatively, we measured the correlation (β) between cell movement and the presence of static and dynamic clusters (Fig. 1b, lower panel). Importantly, the presence of static, but not dynamic, clusters was highly correlated with cell movement (Fig. 1c, $n = 95$).

Close examination of dynamic clusters by TIRF revealed that they not only move between poles but also move across the cell width following a helical path (Fig. 1d). The helicity, characterized by the angle of a helix projected on a plane (φ_a , Fig. 1e), was constant between cells ($\varphi_a = 78^\circ \pm 5^\circ$, $n = 54$; Fig. 1e). In most cases (92%, $n = 54$), the direction of rotation of AglZ–YFP clusters was counterclockwise relative to the direction of movement, denoting a right-handed helical path. Treating the cells with A22, a drug that inhibits MreB polymerization¹⁰, decreased the number of dynamic clusters per cell (Extended Data Fig. 1c, d) but, notably, not their helical movement or directionality (Extended Data Fig. 1e).

¹Laboratoire de Chimie Bactérienne, CNRS-Aix Marseille University UMR7283, Institut de Microbiologie de la Méditerranée, 13009 Marseille, France. ²Centre de Biochimie Structurale, CNRS UMR5048, INSERM U1054, Montpellier University, 29 rue de Navacelles, 34090 Montpellier, France. ³Department of Biology, Indiana University, Bloomington, Indiana 47405, USA. ⁴National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, Maryland 20892, USA. ⁵Interdisciplinary Biochemistry Program, Indiana University, Bloomington, Indiana 47405, USA. ⁶Department of Chemistry, Indiana University, Bloomington, Indiana 47405, USA. ⁷Adhesion and Inflammation laboratory, INSERM U1067, Aix Marseille University UMR7333, 13288 Marseille, France.

*These authors contributed equally to this work.

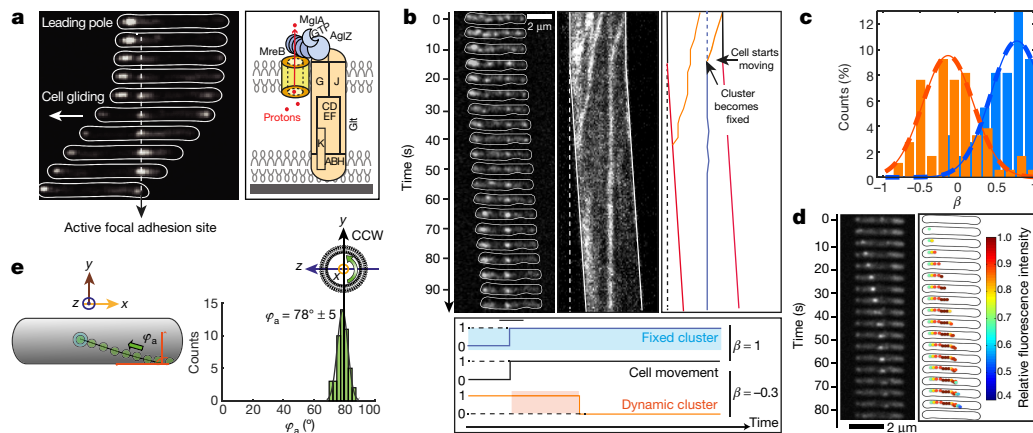


Figure 1 | The *Myxococcus* motility complex moves directionally along a helical path. **a**, Spatial regulation of the *Myxococcus* motility complex. A motile AglZ-YFP-expressing cell showing a complete assembly cycle is shown at 30-s intervals. A current view of an Agl (blue)–Glt (yellow) complex at an FAS is shown^{2,7,19}. Assembly occurs at the leading cell pole following interactions between MglA–GTP, MreB and AglZ (right panel, blue⁸). The position of the Glt proteins is drawn based on published works^{2–5,7,20}. **b**, Immobilization of AglZ-YFP clusters correlates with cell movement. TIRFM of AglZ-YFP in a cell that shifts to motility on a chitosan-treated surface. Images were acquired every 0.5 s. Selected time frames and the corresponding high-resolution kymograph are shown. Two dynamic (orange) clusters are shown. Note that cell movement (indicated by the dashed line showing the initial cell position) is observed

only when a cluster becomes stationary (see orange–blue cluster). Scale bar, 2 μ m. Lower panel: calculation of the correlation coefficient (R) between the presence of a cluster and cell movement. The fixed cluster (blue) is highly correlated with cell movement ($\beta = 1$), while the dynamic cluster (orange) is partially anti-correlated ($\beta = -0.3$). **c**, Distribution of R for fixed (blue) and dynamic clusters (orange), $n = 95$ (six biological replicates). **d**, AglZ-YFP clusters move along helical trajectories. TIRFM at selected time frames of a dynamic AglZ-YFP cluster in a non-motile cell are shown. Scale bar, 2 μ m. **e**, Measurement of the trajectory angle (φ_a) from $n = 54$ (eight biological replicates) single trajectories of dynamic AglZ-YFP clusters (top panels). Histogram of φ_a and Gaussian fit (grey line) are shown. The mean angle is shown with a dashed vertical line and corresponds to counterclockwise trajectories.

Moving cells rotate in a clockwise direction

We reasoned that if propulsion was linked to the counterclockwise trafficking of AglZ-YFP-containing motility complexes, then a gliding cell body should rotate along a similar helical path but of opposite handedness (that is, clockwise; Fig. 2a). To test this idea, we followed the dynamics of fiducial markers (artificial fluorescent D-amino acids, or TADA) fixed to the cell periphery during cell movement (Extended Data Fig. 2a–d). In motile cells, TADA clusters moved from one side

of the cell to the opposite side with angular velocities proportional to the speed of motility (Fig. 2b and Extended Data Fig. 2e), consistent with TADA clusters reporting on the overall rigid-body rotational movement of the cell during propulsion.

To determine the direction of rotation, we directly tracked the 3D positions of TADA clusters during cell movement by inducing astigmatism into the optical detection path¹¹ (see Methods). In agreement with our predictions, TADA clusters rotated in the clockwise direction during cell propulsion (Fig. 2c and Extended Data Fig. 2f; $n = 8$). These observations were confirmed by monitoring the changes in fluorescence intensity produced when TADA clusters moved in and out of the imaging plane during cell movement (see Methods and Extended Data Fig. 2g; $n = 17$). The rotation angle for TADA clusters (φ_T) was constant between cells ($82.6^\circ \pm 2.7^\circ$, $n = 25$, as measured by both methods; Fig. 2d) and closely matched the angle of rotation (φ_a) measured for AglZ-YFP clusters ($78^\circ \pm 5^\circ$). Interestingly, φ_T did not vary significantly with cell speed (Fig. 2d). Together, these findings indicate that anchoring of dynamic AglZ-YFP-containing complexes leads to the clockwise rotation and forward propulsion of the cell.

Functional analysis of the Agl-Glt motility complex

To gain further insight into the molecular mechanism of motility, we genetically dissected the functional groups composing the Agl-Glt machinery (Fig. 1a). Notably, mutations in *agl* and *glt* genes led to aberrant localization patterns and perturbation of AglZ-YFP cluster dynamics (Fig. 3a and Extended Data Fig. 3a). The global effect of each mutation on the localization pattern of AglZ-YFP was evaluated by measuring four observables: the proportion of cells with AglZ-YFP clusters, the mean number of clusters, and the longitudinal and radial distributions of clusters for wild-type and mutant cells (Fig. 3a and Extended Data Fig. 3b). Principal component analysis (PCA) was used to quantitatively characterize the effect of each mutant in the assembly of AglZ-YFP clusters (see Methods). This statistical method allowed us to convert the set of correlated observables into a set of linearly uncorrelated principal components (PCs).

The first three principal components described >87% of the variance (Extended Data Fig. 4a–d). PC1 represented a linear combination of

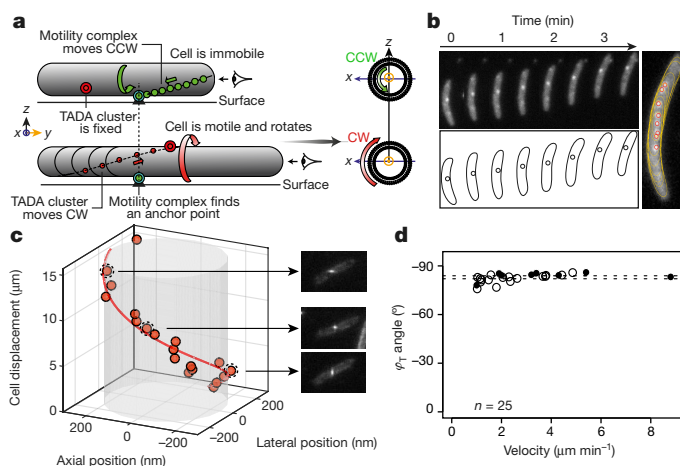
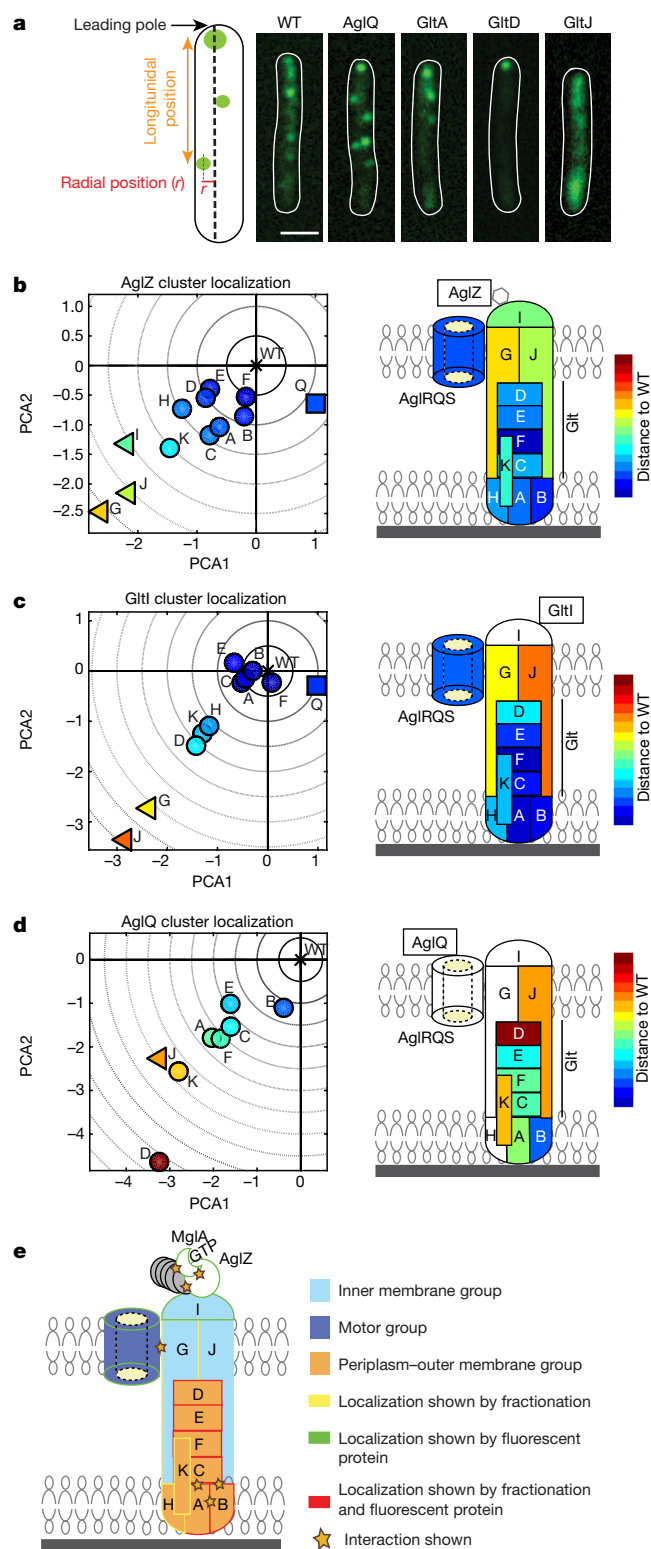


Figure 2 | *Myxococcus* cells rotate during motility. **a**, The helical movement of intracellular motors predicts rotation of the cell during movement. **b**, Rotation of a TADA-bright cluster in a motile cell (30-s intervals) and in a sum-type projection (right). The cartoon representation shows the position of the centroid of the cluster relative to the cell outline. **c**, Three-dimensional positions of a TADA-bright cluster measured by astigmatism (orange dots) illustrating the clockwise rotation of the motile cell ($n = 8$; five biological replicates). Snapshots of the cell are displayed, illustrating the deformation of the microscope point spread function as a function of the axial position. **d**, Rotation angle of TADA-bright clusters (φ_T) as a function of cell velocity from astigmatism (filled circles, $83.9^\circ \pm 2.5^\circ$; $n = 8$, 5 biological replicates) and intensity variations (open circles $82.0^\circ \pm 2.6^\circ$; $n = 17$, three technical replicates).



the number of AglZ-YFP clusters, longitudinal cluster position, and proportion of cells with clusters. PC2 mostly represented the average number of clusters detected per cell, while PC3 was essentially dominated by the mean radial cluster position (Extended Data Fig. 4a–d). A synthetic representation of results can be obtained by plotting the mean position of each mutant in a principal component space where the wild type is arbitrarily positioned at the origin (Fig. 3b; for clarity only PC1 and PC2 are shown but the analysis is computed from PC1–3; Extended Data Fig. 4e). Therefore, for each mutant, the distance to

Figure 3 | Functional architecture of the Agl-Glt machinery. **a**, The localization of AglZ-YFP is affected to different extents in *agl-glt* mutants. The number and position (longitudinal and radial) of clusters was determined in each mutant background and their effect analysed by PCA (see Methods). Scale bar, 1 μ m. **b–d**, The effect of mutations in assembly dynamics of AglZ-YFP, GltI-YFP and AglQ-mCherry plotted in PC coordinates. Triangles represent inner membrane (IM) group proteins, squares represent motor components and circles represent other components (periplasm-outer membrane (OM)). Right panels show schematic representations of the projections of the defined gene groups with respect to their predicted/shown localization from^{2–5,7,20}. Colour code represents distance to the wild type (WT) in PC space. Letters represent each protein subunit of the Agl-Glt complex. **e**, The compilation of PCA (b–d) superimposed to available data^{2–5,7,20} suggests that the motility machinery consists of a molecular motor (dark blue), an inner membrane-cytosolic group (light blue) and a large periplasm-outer membrane group (orange).

the origin and the direction represent the relative effect of a mutation on the number and distribution of AglZ-YFP clusters. Mutations in predicted inner membrane components (GltI, G, J) produced the largest perturbations in the assembly dynamics of AglZ-YFP clusters, whereas mutations in periplasmic outer-membrane group proteins (GltA, C, D, E, H, K) displayed the lowest effect (Fig. 3b). Interestingly, mutations in AglQ (motor) affected AglZ-YFP cluster formation in a qualitatively different manner, represented by a direction in PC space orthogonal to that observed for other mutants.

We refined the functional connections between *agl* and *glt* genes by further investigating the effect of gene deletions on the formation of GltI-YFP⁷ and AglQ-mCherry^{3,7} clusters (GltI and AglQ are putative cytosolic and motor components of the gliding machine, respectively). The overall effects of gene deletions on the number and distribution of GltI-YFP clusters were similar to those observed for AglZ-YFP clusters, consistent with GltI and AglZ belonging to the same functional group (Fig. 3c). Mutations in GltJ and GltG, two predicted inner membrane proteins² that might be in direct contact with GltI, had a large effect on GltI-YFP assembly (Fig. 3c). Interestingly, mutations in GltD, GltH and GltK had a larger effect than did mutations in other proteins in their subgroup (for example, GltA, B, C, E), suggesting that direct protein-protein interactions between these proteins and factors of the GltI subgroup may be needed for GltI-YFP localization (Fig. 3c). Finally, we analysed the effect of deletions in the assembly of motor components (AglQ). Strikingly, deletion of factors in all subgroups led to severe perturbations in the formation of AglQ-mCherry fluorescent clusters (Fig. 3d), suggesting that the motor may require several contacts with different Glt proteins to form functional clusters. However, specific components exhibited differential roles: deletion of GltJ, GltD and GltK had the largest effects, whereas deletion of GltB barely affected the formation of AglQ-mCherry clusters (Fig. 3d). Overall, these data indicate that the motility complex is divided into several distinct functional groups (Fig. 3e).

Motor docks to OM complex during propulsion

To determine the sequence of events leading to the assembly of propulsive complexes, we imaged the dynamic localization of AglZ-YFP (cytosolic-inner membrane group) simultaneously with that of proteins belonging to each of the other functional subgroups: AglQ (motor), GltD (periplasmic), or GltC-mCherry (periplasm-outer membrane complex) (Fig. 4a, c, e). Under conventional epifluorescence microscopy, cytosolic (AglZ) and motor components (AglQ) in motile cells (on 1.5% agar) displayed a strong ability to form clusters, in contrast to the dispersed localization of periplasmic and outer membrane complex components (GltC, D) to the cell periphery (Extended Data Fig. 5a). Thus, we used TIRFM to image the dynamic localization of these factors in motile cells adhered to chitosan-coated glass. AglQ-mCherry and AglZ-YFP co-localized in both stationary and mobile clusters, indicating that these proteins form a stable

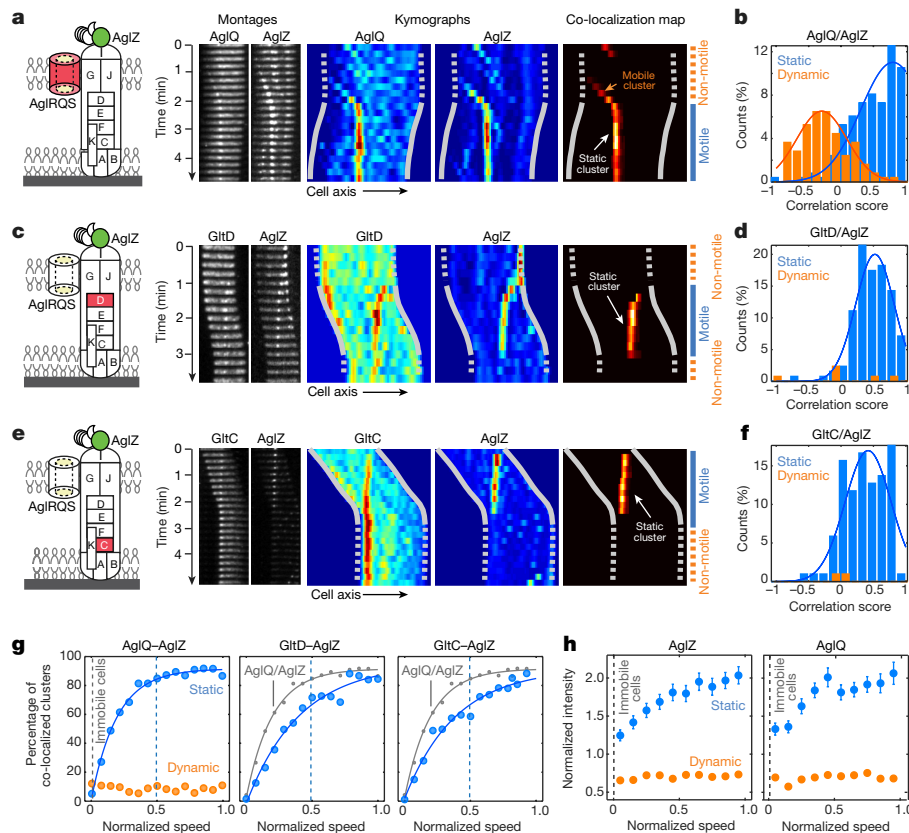


Figure 4 | Dynamic interactions between the intracellular and outer membrane complexes generate propulsion. **a**, Simultaneous dynamics of AglZ-YFP and AglQ-mCherry and correlation to the motility phases. TIRFM time-lapse images and corresponding kymographs are shown for individual fluorescent fusions and in a heat map showing the computed co-localization scores. **b**, AglZ and AglQ co-localize in the inner membrane trafficking complex ($n = 156$ cells; two biological replicates). Correlation scores reflect the correlation between cell movement and static co-localized clusters. **c**, **d**, Simultaneous dynamics of AglZ-YFP and GltD-mCherry and correlation to the motility phase ($n = 121$ cells; seven biological replicates). Legend as in **a**, **b**, **e**, **f**, Simultaneous dynamics of AglZ-YFP and GltC-mCherry and correlation to the motility phase

($n = 100$ cells; three biological replicates). Legend as in **a**, **b**, **g**, Correlation between cell velocity and number of active complexes. For a given value of the cell speed (V), the percentage of co-localized clusters represents the proportion of events in which a cell was moving at speed V and an active cluster containing both AglZ and AglQ, GltD or GltC was detected. The percentage of co-localization is shown as a function of normalized cell speed (where 1 is the maximum cell speed). **h**, Correlation between cell velocity and intensity of AglZ and AglQ clusters. Intensity was normalized with respect to the average intensity measured for dynamic clusters and shown as a function of normalized cell speed. Error bars represent s.e.m. calculated according to the total number of cells analysed ($n = 114$ cells for AglZ, $n = 100$ for AglQ).

complex in the bacterial inner membrane (Fig. 4a; see also Methods and Extended Data Figs 5b, c, 6). To investigate whether clusters containing both AglZ and AglQ were linked to cell motility, we calculated for each cell the correlation between cell movement and co-localization and computed the distributions for both stationary and mobile clusters (Fig. 4b; see also Methods and Extended Data Fig. 7). The presence of stationary clusters was highly correlated to cell movement, whereas that of mobile clusters was not, consistent with AglZ and AglQ being part of the trafficking internal complex (Fig. 4b). By contrast, GltD-mCherry and GltC-mCherry co-localized with AglZ-YFP only in stationary clusters (Fig. 4c, e and Extended Data Fig. 5b, c) and co-localization of stationary AglZ-GltC/D clusters was highly correlated to cell movement (Fig. 4d, f). These data strongly suggest that cytosolic-outer membrane and motor complexes assemble together in a mobile unit that requires a physical connection to periplasmic and outer membrane components to form stationary clusters that impart cell movement.

To investigate this hypothesis further, we determined the correlation between cell speed and the proportion of clusters containing both AglZ and AglQ, GltC or GltD (Fig. 4g). Dynamic clusters containing AglZ and AglQ were observed in motile cells, but their number was independent of cell speed, indicating that they are not propulsive by themselves (Fig. 4g). On the contrary, the proportion of AglZ-AglQ clusters that were stationary increased rapidly with cell speed (Fig. 4g). The number of AglZ-AglQ clusters reached a maximum at $\sim 50\%$ of

the cell maximum speed ($V_{1/2}$), suggesting that this number is not a limiting factor (Fig. 4g). In fact, the recruitment of GltD and GltC to FASs could be a limiting step because a substantial number of static AglZ clusters lack GltD and, especially, GltC (Extended Data Fig. 5b, c). Consistent with this finding, the number of stationary AglZ-GltD and AglZ-GltC clusters was only $\sim 60\%$ of the maximum at $V_{1/2}$ and full speed was reached only when the percentage of GltD-GltC stationary clusters recruited to FASs saturated (Fig. 4g). Thus, the recruitment of GltD and GltC to FASs is required for propulsion, probably because these proteins belong to complexes that link the motor to the external surface.

To test further whether the stoichiometry of Agl-Glt components regulates the activity of FASs, we measured how the mean fluorescence intensity of stationary and mobile clusters changed with normalized cell speed (Fig. 4h and see Methods). Cell speed increased with the number of AglZ and AglQ subunits accumulating in stationary clusters (Fig. 4h). By contrast, the proportion of these subunits in mobile clusters was systematically lower and uncorrelated with cell speed (Fig. 4h). All in all, these results show that stationary focal adhesion complexes form as a result of the transient recruitment of periplasmic and outer membrane proteins by the mobile PMF-driven inner membrane complex. Importantly, the activity of stationary complexes is regulated at two levels: by the number of Agl motor units (force generation) and by the number of local contacts with periplasmic-outer membrane components (transmission).

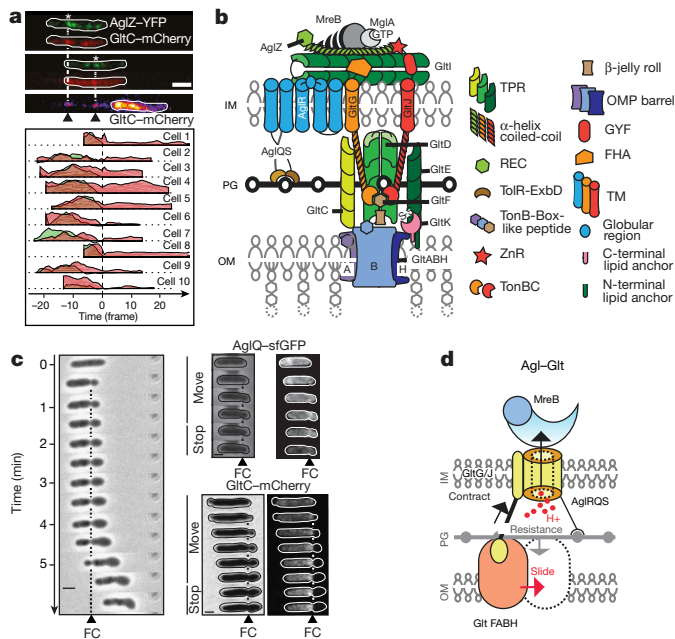


Figure 5 | Cyclic contacts between the inner membrane motor and outer membrane adhesins drive propulsion. **a**, GltC-mCherry is released by gliding cells at focal adhesion complexes. Top, TIRFM snapshots for a representative cell expressing both GltC-mCherry (red) and AglZ-YFP (green). The positions of the GltC clusters on the surface coincide with the positions of focal adhesion complexes (white asterisk). Time frames, 15 s. Scale bar, 2 μ m. Bottom, variation in intensity of GltC-mCherry (red) and AglZ-YFP (green) as a function of time before (negative time) and after (positive time) the cell moved away from the focal adhesion position ($n = 10$; two technical replicates). **b**, Predicted domain architecture of the Agl-Glt machinery based on bioinformatics prediction, sequence analysis and previous literature. The different proteins of the complex are represented on the basis of their domain structures from bioinformatics predictions (Supplementary Table 1). Each protein is represented as a single copy in the complex. **c**, A fixed contractile zone is observed in motile sporulating cells where peptidoglycan is profoundly remodelled. Motile cells in the early phases of sporulation are shown at 1-min intervals. AglQ-sfGFP and GltC-mCherry are specifically enriched at the constriction site. Representative wild-type ($n = 10$), AglQ-sfGFP ($n = 2$, two technical replicates) and GltC-mCherry ($n = 4$, two technical replicates) cells are shown, obtained from two biological replicates. **d**, Possible mechanism of propulsion. The structure of the Agl-Glt machinery is simplified to its core components for clarity. The proton flow through a peptidoglycan-bound TolQR-type channel (yellow) is proposed to energize cyclic interactions between a flexible inner membrane-anchored periplasmic protein (GltG/J, black and yellow) and TonB-box proteins in the outer membrane (orange). Combined with the rigid anchoring to peptidoglycan and link with MreB, this activity would push the outer membrane protein laterally (red arrow) because peptidoglycan counteracts the exerted traction force (grey arrow). The protein stoichiometries are not known and it is possible that the complex contains several coordinated legs, facilitating the processivity and directionality of the movements.

Adhesion and cyclic OM-IM contacts observed at FAS

A clear prediction of this model is that the motility complex should require direct adhesive contacts with the underlying surface to propel the cell. Consistent with this prediction, reflection interference contrast microscopy (RICM) revealed that motile cells are uniformly in contact with the substratum and high RICM densities were correlated with the position of AglZ-YFP clusters (Extended Fig. 5d). Notably, vesicles containing only outer membrane materials and specific outer membrane Glt proteins (GltC-mCherry) were deposited at sites coinciding with the positions of focal adhesion complexes in the wake of motile cells (Fig. 5a and Extended Data Fig. 5e, f). Overall, these data suggest that adhesions involve strong, intimate contacts between the surface and the periplasm-outer membrane complex.

To gain molecular insight into the mechanisms involved in force generation and surface adhesion, we predicted a protein-domain structure of the motility complex using bioinformatics approaches (Fig. 5b and Extended Data Figs 8, 9). Remarkably, both AglQ and AglS contain predicted TolR-like PG-binding motifs¹² and AglR, a TolQ homologue interacts with GltG² a TolA/TonB-like protein (Extended Data Fig. 8a-b). GltG and GltJ are similar modular proteins, specifically sharing a single transmembrane helix, a periplasmic helical domain, and TonBC motifs¹³. Interactions of GltGJ with OM components could occur between their TonBC domains and potential TonB-box-carrying proteins -GltF in the periplasm (Extended Data Fig. 8a) and GltAB, two predicted porin-like β -barrel proteins in the OM (Extended Data Fig. 9a-b).

Thus, peptidoglycan-anchored Agl motor units may act as staters, pushing against adhesive outer membrane complexes through peptidoglycan to generate propulsive forces. We reasoned that these dynamic interactions might be revealed under conditions in which the rigidity of the cell envelope is reduced, for example in cells undergoing sporulation where rapid peptidoglycan remodelling leads to the formation of round cells¹⁴. Indeed, the induction of sporulation rapidly led to the formation of balloon-shaped cells (Fig. 5c). Notably, a small fraction of cells (<1%) entering sporulation were still motile and formed conspicuous constrictions that, similar to focal adhesion complexes, remained fixed relative to the surface (Fig. 5c). These constrictions probably resulted from the activity of the Agl-Glt machinery because: (i) they formed only in motile sporulating cells (<1% of the sporulating cell population); and (ii) both AglQ-sfGFP (inner membrane) and GltC-mCherry (outer membrane) were enriched at constriction sites (Fig. 5c). Thus, in sporulating cells where the structure of peptidoglycan is different (peptidoglycan is not detected in mature spores¹⁵), dynamic Agl-Glt-driven physical contacts between the outer membrane and the inner membrane are unmasked.

Discussion

On the basis of our results, we propose that the Agl motor and associated inner membrane proteins move directionally by cyclic interactions with factors in the periplasmic-outer membrane complex (Fig. 5d). Analogous to Tol-Exb systems^{13,16}, these steps would occur by PMF-driven conformational changes in the Glt TolA/TonB-like proteins (GltG and/or GltJ; Fig. 5b), the flexible domains of which might extend and retract through the peptidoglycan layer (Fig. 5d). Peptidoglycan itself could act as both a transient anchor point, as it becomes bound by AglQ/S via a TolR-like peptidoglycan-binding motif, and a guiding factor, as it opposes contractions to favour lateral movements (Fig. 5d).

The current study does not resolve the relative stoichiometries of the Agl and Glt proteins at focal adhesions, but the data show that the activity of focal adhesion complexes is subject to regulation and contains variable stoichiometries of Agl motor components and connections with the underlying surface. Thus, it is possible that several 'legs' operate coordinately at these sites. The directionality of motility complexes is remarkably consistent between cells, suggesting that a core cellular structure is involved; peptidoglycan is an attractive candidate because the glycan strands are proposed to have a global right-handed helical ordering that could serve as tracks to guide the motility complex¹⁷. At FASs, the interaction with the surface appears to be strong, implying the existence of one or more specific adhesion molecule(s); consequently, there is likely to be a relief mechanism similar to that found in gliding parasites, where the major adhesion molecule is removed from the motility complex by specific proteolysis¹⁸. In addition, our model makes several important predictions that will help future studies, such as the systematic exploration of protein interactions in the system to attain a molecular understanding of the motility mechanism.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 21 March; accepted 28 September 2016.

Published online 5 October 2016.

- Islam, S. T. & Mignot, T. The mysterious nature of bacterial surface (gliding) motility: A focal adhesion-based mechanism in *Myxococcus xanthus*. *Semin. Cell Dev. Biol.* **46**, 143–154 (2015).
- Luciano, J. *et al.* Emergence and modular evolution of a novel motility machinery in bacteria. *PLoS Genet.* **7**, e1002268 (2011).
- Sun, M., Wartel, M., Cascales, E., Shaevitz, J. W. & Mignot, T. Motor-driven intracellular transport powers bacterial gliding motility. *Proc. Natl Acad. Sci. USA* **108**, 7559–7564 (2011).
- Jakobczak, B., Keilberg, D., Wuichet, K. & Søgaard-Andersen, L. Contact- and protein transfer-dependent stimulation of assembly of the gliding motility machinery in *Myxococcus xanthus*. *PLoS Genet.* **11**, e1005341 (2015).
- Nan, B. *et al.* Myxobacteria gliding motility requires cytoskeleton rotation powered by proton motive force. *Proc. Natl Acad. Sci. USA* **108**, 2498–2503 (2011).
- Balagam, R. *et al.* *Myxococcus xanthus* gliding motors are elastically coupled to the substrate as predicted by the focal adhesion model of gliding motility. *PLOS Comput. Biol.* **10**, e1003619 (2014).
- Treuner-Lange, A. *et al.* The small G-protein MglA connects to the MreB actin cytoskeleton at bacterial focal adhesions. *J. Cell Biol.* **210**, 243–256 (2015).
- Nan, B. *et al.* Flagella stator homologs function as motors for myxobacterial gliding motility by moving in helical trajectories. *Proc. Natl Acad. Sci. USA* **110**, E1508–E1513 (2013).
- Kaiser, D. & Warrick, H. Transmission of a signal that synchronizes cell movements in swarms of *Myxococcus xanthus*. *Proc. Natl Acad. Sci. USA* **111**, 13105–13110 (2014).
- Bean, G. J. *et al.* A22 disrupts the bacterial actin cytoskeleton by directly binding and inducing a low-affinity state in MreB. *Biochemistry* **48**, 4852–4857 (2009).
- Huang, B., Wang, W., Bates, M. & Zhuang, X. Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science* **319**, 810–813 (2008).
- Wojdyla, J. A. *et al.* Structure and function of the *Escherichia coli* Tol-Pal stator protein TolR. *J. Biol. Chem.* **290**, 26675–26687 (2015).
- Gresock, M. G., Kastead, K. A. & Postle, K. From homodimer to heterodimer and back: elucidating the TonB energy transduction cycle. *J. Bacteriol.* **197**, 3433–3445 (2015).
- Wartel, M. *et al.* A versatile class of cell surface directional motors gives rise to gliding motility and sporulation in *Myxococcus xanthus*. *PLoS Biol.* **11**, e1001728 (2013).
- Bui, N. K. *et al.* The peptidoglycan sacculus of *Myxococcus xanthus* has unusual structural features and is degraded during glycerol-induced myxospore development. *J. Bacteriol.* **191**, 494–505 (2009).
- Cascales, E., Gavioli, M., Sturgis, J. N. & Lloubès, R. Proton motive force drives the interaction of the inner membrane TolA and outer membrane pal proteins in *Escherichia coli*. *Mol. Microbiol.* **38**, 904–915 (2000).
- Wang, S., Furchtgott, L., Huang, K. C. & Shaevitz, J. W. Helical insertion of peptidoglycan produces chiral ordering of the bacterial cell wall. *Proc. Natl Acad. Sci. USA* **109**, E595–E604 (2012).
- Ejigiri, I. *et al.* Shedding of TRAP by a rhomboid protease from the malaria sporozoite surface is essential for gliding motility and sporozoite infectivity. *PLoS Pathog.* **8**, e1002725 (2012).
- den Blaauwen, T., de Pedro, M. A., Nguyen-Distèche, M. & Ayala, J. A. Morphogenesis of rod-shaped sacculi. *FEMS Microbiol. Rev.* **32**, 321–344 (2008).
- Nan, B., Mauriello, E. M. F., Sun, I.-H., Wong, A. & Zusman, D. R. A multi-protein complex from *Myxococcus xanthus* required for bacterial gliding motility. *Mol. Microbiol.* **76**, 1539–1554 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank R. Mercier for suggestions and critical reading of the manuscript, A. Valeri for help with PCA analysis, C. Fiche for help with image analysis and Y. Denis (IMM transcriptomics platform) for help with quantitative PCR and M. Wartel for preliminary experiments. L.M.F. was partially funded by Fondation ARC, L.E. was supported by the Fondation pour la Recherche Médicale (FRM DGE2010221257), S.T.I. was supported by a fellowship from the Canadian Institutes of Health Research and the AMIDEX program of Aix-Marseille Université, and A.L. and V.A. were supported by the IRP funds of the NIH, NIH grant GM113172 to M.S.V.N. and Y.V.B. and NIH grant GM51986 to Y.V.B. Research in the Mignot laboratory was supported by European Research Council grant DOME-261105 and a Bettencourt-Schueller “Coup d’élan pour la recherche Française 2011”. Research in the Nollmann laboratory was supported by ANR grants IBM (ANR-14-CE09-0025-01), HiResBacs (ANR-15-CE11-0023), and European Research Council grant smlInsulator-260787. We acknowledge support from France-BioImaging (ANR-10-INBS-04, “Investments for the future”), and Imagine Optic.

Author Contributions L.M.F., J.-B.F., L.E., M.N. and T.M. conceived the experiments and analysed the data with help from A.D. for TADA experiment analysis. L.M.F. and J.-B.F. performed most of the experiments, including TIRF assays, astigmatism and strain construction. A.D. and S.L. helped with TADA experiments. J.L., M.S., and S.L. constructed strains. S.T.I., J.T. and O.T. performed RCM studies. V.A. and A.L. performed bioinformatics analysis. E.K., M.S.V.N. and Y.V.B. provided TADA. M.N. and T.M. wrote the paper with help from L.M.F., J.-B.F. and L.E.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to T.M. (tmignot@imm.cnrs.fr) or M.N. (marcelo.nollmann@cbs.cnrs.fr).

Reviewer Information Nature thanks J. Armitage, P. Kukura and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

Bacterial strains, plasmids, and growth. Strains, primers, and plasmids are listed in Supplementary Tables 2 and 3. See Supplementary Tables 2 and 4 for strains and their mode of construction. *M. xanthus* strains were grown at 32 °C in caseitone yeast extract (CYE)-rich medium as previously described²¹. Plasmids were introduced into *M. xanthus* by electroporation. Mutants and transformants were obtained by homologous recombination based on a previously reported method²¹. All fusions were expressed by gene replacement at the endogenous loci allowing expression from the natural promoters. Quantitative PCR experiments confirmed that expression was similar to wild-type levels (expression ratios varied between 1 and 2 compared to wild-type levels). Three types of AglQ fusion were used throughout the study: TIRF experiments used AglQ–mCherry expressed from the endogenous locus in place of the wild-type gene. For technical reasons, the PCA analysis and expression of AgQ–sGFP used fusions expressed after ectopic integration of the gene of interest at the Mx8-phage attachment site in a $\Delta aglQ$ deletion background (Supplementary Table 4). All three strains were indistinguishable in terms of motility and expression patterns. *Escherichia coli* cells were grown under standard laboratory conditions in Luria–Bertani broth supplemented with antibiotics, if necessary.

Motility assays on agar surfaces. For standard phase-contrast and fluorescence microscopy, cells from exponentially growing cultures were transferred to a 1.5% agar pad with TPM buffer (10 mM Tris-HCl, pH 7.6, 8 mM MgSO₄, and 1 mM KH₂PO₄) on a glass slide and covered with a coverslip. Imaging was performed in a temperature-adjusted microscope chamber at 32 °C. To test the function of peptidoglycan during motility, sporulation was induced by adding 5% glycerol directly into the agar pads as previously described¹⁴. This treatment induces rapid and controlled degradation of the *Myxococcus* peptidoglycan leading to the formation of viable spheroplasts¹⁴. Cells were imaged 15 min after deposition on a glycerol pad in order to image their motility while converting to spheroplasts.

Microfluidics on chitosan-coated glass slides. *M. xanthus* cells were immobilized on a chitosan-coated surface, as described previously²². In brief, custom-built polydimethylsiloxane microfluidic glass chambers were coated with chitosan solution and washed after 30 min. Chambers were further rinsed with 1 ml TPM buffer (10 mM Tris-HCl, pH 7.6, 8 mM MgSO₄, 1 mM KH₂PO₄). Subsequently, 1 ml of an exponentially growing culture was injected into the chamber and left for 30 min without flow. Unattached cells were removed by rinsing with 1 ml TPM buffer by manual injection and time-lapse microscopy on attached cells was performed. When needed, A22 (Merck Millipore) was injected manually at indicated concentrations a few minutes after the cells were confirmed to be motile.

Labelling cells with fluorescent D-amino acids. Lyophilized TADA (molecular mass = 381.2 g/mol) powder was re-suspended in DMSO at 150 mM and conserved at –20 °C. The labelling was performed for 2 h at 32 °C, using 2 μ l TADA solution for 1 ml cell culture (OD₆₀₀ = 0.5) in the presence of 1 M NaCl for a duration of 2 h. The sample was then washed four times with 1 ml TPM buffer and finally resuspended at OD₆₀₀ = 5 before being transferred to an agar pad.

Bioinformatics analyses. Iterative sequence profile searches were performed using the PSI-BLAST²³ and JACKHMMER²⁴ programs run against the non-redundant protein database of National Center for Biotechnology Information (NCBI). Similarity-based clustering for both classification and culling of nearly identical sequences was performed using the BLASTCLUST program (<ftp://ftp.ncbi.nlm.nih.gov/blast/documents/blastclust.html>). The HHpred program was used for profile–profile comparisons²⁵. Structure similarity searches were performed using the DaliLite program²⁶. Multiple sequence alignments were built by the Kalign and PCMA programs^{27,28}, followed by manual adjustments on the basis of profile–profile and structural alignments. Secondary structures were predicted using the JPred program²⁹. For previously known domains, the Pfam database was used as a guide³⁰, and the profiles were augmented by addition of newly detected divergent members that were not detected by the Pfam models. Clustering with BLASTCLUST followed by multiple sequence alignment and further sequence profile searches were used to identify other domains that were not present in the Pfam database. Signal peptides and transmembrane segments were detected using the TMHMM and Phobius programs^{31,32}. Contextual information from prokaryotic gene neighbourhoods was retrieved using custom Perl scripts that extract the upstream and downstream genes of the query gene, along with their orientation, from GenBank files. A combination BLASTCLUST and sequence profile searches were then used to cluster the proteins to identify conserved gene-neighbourhoods. Structural visualization and manipulations were performed using the PyMol (<http://www.pymol.org>) program. The in-house TASS package, which comprises a collection of Perl scripts, was used to automate aspects of large-scale analysis of sequences, structures and genome context.

Epifluorescence microscopy. For regular epifluorescence, GFP or mCherry fluorochromes were visualized at 32 °C using a temperature-controlled TE2000-

E-PFS microscope (Nikon) with a 100 \times NA 1.3 (PhC) objective and a CoolSNAP HQ2 camera (Photometrics). All fluorescence images were acquired with appropriate filters with a minimal exposure time to minimize bleaching and phototoxicity effects. Images were recorded with Metamorph software (Molecular Devices).

TIRFM. TIRFM was performed with an inverted microscope (Axio Observer A1, Zeiss) equipped with a 100 \times Plan-Achromat oil-immersion objective (NA = 1.46) mounted on a closed-loop piezoelectric stage (PIFOC, Physik Instrumente). Two lasers with excitation wavelengths of 488 nm (OBIS 488LS, Coherent USA) and 561 nm (Sapphire 561LP, Coherent USA) were used for YFP and mCherry imaging, respectively. Laser beams were combined and collimated by a series of dichroic mirrors and achromatic lenses, individually controlled by an acousto-optic tunable filter (AOTF, AA Opto-electronic) and focused onto the back focal plane of the objective through the rear port of the microscope. A translation stage was used to shift the positions of the two beams with respect to the objective, enabling an easy permutation between Epi and TIRF imaging (Applied Scientific Instrumentation). The fluorescence emission signal was collected by the objective lens, separated from the excitation wavelengths through a four-band dichroic mirror and filtered using bandpass filters inserted in a high-speed motorized filter wheel (Chroma Technology). The filtered emission signal was then imaged onto an emCCD camera (Andor Ixon 897) through relay lenses allowing an effective pixel size of 105 nm. Along a separate path, a 785-nm infrared laser beam was focused on the back focal plane of the objective and reached the glass–sample interface in total internal reflection conditions. The reflected infrared beam was imaged by a CMOS camera (Thorlabs Inc.) and its position calculated during live acquisition. The information was fed-back to a z-positioning piezo stage through PID software to correct for any change in the objective–sample distance. This active autofocus system locks the focal plane position with a precision of ± 20 nm over hours. All acquisition software controlling lasers, filter wheel, translation stages and cameras were homemade using LabView 2012 (National Instruments). For time-lapse imaging, a high-speed motorized filter wheel was used to sequentially image YFP and mCherry channels with a switching time of less than 200 ms. Typically, 5–10 images were acquired at 20 Hz for each channel and the process was repeated 45–50 times every 15–30 s, depending on the cell speed. For real-time imaging, 500 images were taken at 20 Hz in the YFP channel. Laser intensity was optimized to get the best signal-to-noise ratio (SNR) while limiting photobleaching and phototoxicity. In order to compensate for chromatic aberrations, each channel was assigned a specific setpoint for the autofocus depending on the position of the focal planes for YFP and mCherry fluorescence signals.

Astigmatism for 3D time-lapse experiments. First, analysis of TADA-bright clusters by epifluorescence showed that they are diffraction limited and circular in shape, making their 3D tracking by astigmatism possible. 3D imaging was performed as previously described³³. Briefly, to obtain 3D imaging conditions, a corrective MicAO 3D-SR system (Imagine Optic) was inserted into the emission pathway between a modified Nikon Eclipse Ti-S inverted microscope and an EMCCD camera (Andor Ixon 897). A MicAO 3D-SR adaptive optics device was used to correct the microscope point spread function (PSF) for optical aberrations introduced in the imaging path and to optimize the photon budget. For 3D detection, subtle changes in astigmatism were further added in order to break the axial symmetry of the PSF and allow an estimation of the axial position of diffraction-limited fluorescent objects. Time-lapse imaging of TADA clusters was performed after transferring to an agar pad freshly labelled cells mixed with 100-nm fluorescent beads (TetraSpeck, Thermo Scientific), used as fiducial and calibration markers. Every 15–30 s, two series of 5–10 successive images were simultaneously acquired at 20 Hz: one under continuous epifluorescence illumination with a 561-nm read-out laser to detect the fluorescence of the TADA clusters (Sapphire 561LP, 150 mW, Coherent) and the second using a brightfield illumination to get an image of the cells. Then, a calibration was performed on the same field of view by imaging single fluorescent beads while scanning the sample along the optical axis (z) by steps of 50 nm.

Detection and quantification of TADA clusters is described below. In total, data from eight cells were successfully analysed (five biological replicates). However, the behaviour described was clearly observed on numerous occurrences ($n = 10$), though the data were not of high enough quality (low SNR, cell to cell contact, cell moving too quickly or out of the observation field and so on) to obtain long trajectories.

RICM. For label-free imaging of the underside of cells in contact with the substratum, a modified form of reflection imaging for bacteria was employed³⁴. RICM was performed on chitosan-coated microfluidic chambers. Channels were seeded with $\Delta pilA$ cells expressing AglZ–YFP resuspended in TPM buffer with 1 mM CaCl₂ (OD₆₀₀ = 0.5) for 10 min, then washed with the same buffer. Images were obtained at 32 °C on a Zeiss Axiovert 200 inverted microscope

with adjustable aperture and field stops. For imaging, an RICM objective (Zeiss Neofluar 63/1.25 antifix) and a differential interference contrast objective (Plan-Apochromat 63 × 1.40 oil) were used for crossed-polarized light and fluorescence images, respectively. For RICM, cells were illuminated through a 546 ± 12 nm narrow bandpass filter with a mercury lamp (X-cite 120Q lamp) for 20 ms. For fluorescence, samples were excited with a laser at 488 nm for 1,000 ms. Images were captured at 10-s intervals and combined in ImageJ.

Fluorescence recovery after photobleaching (FRAP). FRAP was performed with a 488-nm laser mounted directly on the TE2000-E-PFS microscope (Nikon), allowing us to focus a micrometre-radius laser beam with micrometre precision. FRAP acquisitions were performed with a home-developed macro under Metamorph.

Cluster image analysis. Trajectories of fluorescent clusters were obtained from the coordinates calculated in a cylindrical reference system based on the shape of the cell. This step was performed automatically and verified manually using the MicrobeJ plugin (<http://www.indiana.edu/~microbej/>) in FIJI/ImageJ³⁵. These coordinates were used to calculate angle, speeds, mean square displacements and angles using R software. To calculate the slope of helicoidal trajectories, the unwrapped coordinates were fitted with a linear regression model and φ angles were determined as the arctangent of the line coefficient (Extended Data Fig. 2e). In TIRFM, owing to the narrow depth of focus, different fluorescent objects (AglZ-YFP) were treated as if they were in the same plane. This simplified the method of calculating the trajectory angles. A z-stack projection of maxima was applied (FIJI). Some cells in the projection image showed clear linear alignments of clusters crossing the cell body. The angle between the linear clusters trajectory and the major axis of the cell was manually measured with the angle tool of the FIJI software.

Principal component analysis. The number of clusters and the relative positions of AglZ-YFP, GltI-YFP and AglQ-mCherry in *agl:glt* mutant backgrounds were obtained by combining the mask of cells obtained from phase contrast images and the fast Fourier transform (FFT)-filtered fluorescent image (FIJI, Extended Data Fig. 3b). In brief, the phase-contrast image of rod-shaped cells provides a mask of cell bodies and yields morphological parameters and definition of the longitudinal axis. Following a straightening operation, this axis can be used as a reference frame for cluster localization. Because the internal clusters are of weak intensity, the images were denoised by applying a neutral density filter and background subtraction. The presence of fluorescent clusters was systematically verified on unprocessed images to ensure that the procedures did not generate artefactual signals. The following list provides the number of cells analysed for each condition resulting from six technical replicates. For the AglZ-YFP reporter: wild-type, 67; Δ gltA, 146; Δ gltB, 407; Δ gltC, 224; Δ gltD, 139; Δ gltE, 168; Δ gltF, 136; Δ gltG, 323; Δ gltH, 114; Δ gltI, 193; Δ gltJ, 174; Δ gltK, 194; Δ gltQ, 407. For the GltI-YFP reporter: wild-type, 187; Δ gltA, 688; Δ gltB, 292; Δ gltC, 321; Δ gltD, 133; Δ gltE, 296; Δ gltF, 114; Δ gltG, 64; Δ gltH, 261; Δ gltI, 359; Δ gltK, 694; Δ gltQ, 187. For the AglQ-mCherry reporter: wild-type, 163; Δ gltA, 124; Δ gltB, 723; Δ gltC, 266; Δ gltD, 201; Δ gltE, 399; Δ gltF, 196; Δ gltG, 189; Δ gltI, 511; Δ gltK, 197.

Two-colour TIRFM. Image analysis was performed using Matlab 2015 (MathWorks). For each experiment, only cells displaying gliding displacement were analysed. For each selected bacterium, a temporal RGB image was calculated using mCherry fluorescence images and the cell trajectory path was manually drawn (Extended Data Fig. 6a). A montage and its associated kymograph were then calculated for both YFP and mCherry channels by straightening and re-slicing each time-lapse image along that path (Extended Data Fig. 6b). For the kymograph, the best contrast was obtained by averaging the intensity of the three brightest pixels (over a total of 13 pixels) for each slice. To further improve the contrast of the kymographs and highlight the presence of fluorescent clusters, a denoising algorithm (modified from ref. 36) was applied (Extended Data Fig. 6c). Photobleaching was also quantified for each channel based on the variation of non-specific fluorescence signal measured in the cells over time. A single exponential model was used to fit the variation of intensity and correct the kymograph intensity accordingly. Small shifts induced by chromatic aberrations between the two channels were corrected as well by realigning the two kymographs using an image cross-correlation algorithm (precision of 1 pixel, 105 nm). Finally, for the co-localization calculation, the intensities of both images were standardized in order to compare the fluorescent signals from both YFP and mCherry channels.

Cell tracking in kymographs. To calculate the cell displacement over time, the kymograph displaying the best contrast was first interpolated: the number of pixels was increased by twofold in each dimension and the intensity of each pixel recalculated using a cubic interpolation algorithm. An iterative edge-detection algorithm based on the Canny method was then applied to the interpolated image³⁷. Typical values for the standard deviation were between 3 and 5 pixels. By iteratively changing the Canny detection threshold, the algorithm converged towards two

different edges highlighting the movement of both cell poles during the time-lapse. Errors in the edge calculation were sometimes observed when the lagging pole of the cell was not properly attached to the chitosan-treated surface and/or when other bacteria were in contact with the selected cell. In that case, a manual correction was performed in order to remove inconsistent points and rectify the position of the edges.

Often, the two edges were not exactly identical owing to imprecision in the localization of poles. Indeed, we could observe that the leading pole of the cell was usually very well defined (highest fluorescence intensity) resulting in a very precise calculation of its localization over time. On the contrary, the lagging pole was often poorly defined due to lower fluorescence intensity and/or weaker adhesion to the chitosan-treated surface. Therefore, in order to quantify the cell displacement during the time-lapse, the edge defined with the highest precision was selected. The speed was also estimated from the first derivative of the displacement and analysed in order to automatically identify when the cell was mobile (gliding) during the time-lapse acquisition. To do so, for each experiment, a velocity threshold was empirically estimated by analysing the movement of a small subset of cells. Then, for each time-point, bacteria were classified as either mobile or immobile depending on whether their velocity surpassed the velocity threshold. Typical values for the threshold lay between 0.05 and 0.2 $\mu\text{m}/\text{min}$, depending on experimental drift and noise (Extended Data Fig. 6d). In average, cell speed on the chitosan-coated surface was $0.5 \pm 0.5 \mu\text{m}/\text{min}$.

Co-localization. Co-localization between YFP-tagged and mCherry-tagged proteins was calculated in three steps (Extended Data Fig. 7a). First, the two montages were separately analysed in order to localize the fluorescent clusters in each channel. For each montage, its normalized cross-correlation with a Gaussian spot (s.d. of 1 pixel) was computed. Fluorescent aggregates were then delimited by thresholding the image (typical value 0.55–0.75) and grouping the selected pixels in separate clusters based on their connectivity. Clusters composed of fewer than 4 pixels were systematically discarded.

In a single cell, many different fluorescent clusters $C_n = 1, 2, \dots$ could be simultaneously detected, each with a different trajectory. As the same fluorescent cluster C_n could be detected in several successive images, localizations were grouped by manually stitching together its detections ($C_n^i, C_n^{i+1}, \dots, C_n^{i+k}$) on the montage. Therefore, each cluster $C_n = 1, 2, \dots$ was now defined by all its localizations during the time-lapse acquisition. By analysing their positions over time, each observation of the same cluster was labelled as either stationary or dynamic. Finally, it is important to point out that clusters detected at the poles were not taken into account for the analysis. The pixel size in our experiments was 105 nm; thus, the precision of detection of the maximum of intensity of clusters from kymographs was approximately ~ 100 nm.

Next, for each montage, the localizations of the clusters were projected onto the associated normalized kymograph. A protein detection map was then calculated by keeping only the intensities of the pixels that were part of a cluster (Extended Data Fig. 7b). The intensity of all other pixels was set to zero. By superimposing the trajectories of the two cell poles, it was then possible to study how the detection of protein clusters was correlated with cell displacement. Each row of the map corresponded to a snapshot of the protein localization and intensity within the cell at a given time of the experiment. Finally, the co-localization map was calculated by multiplying the two detection maps together (Extended Data Fig. 7c). From the distinction performed earlier between stationary and dynamic clusters, we could measure how the co-localization intensity varied over time for both types of clusters. Two curves were therefore calculated by summing along each row the intensity of the pixels associated to either dynamic or stationary clusters. In the end, the total co-localization signal was obtained by summing together the two curves.

Correlation between cell movement and co-localization. For each cell, correlation between movement and protein co-localization was calculated, taking static and dynamic clusters into account separately (Extended Data Fig. 7d, e). For each time-point, the correlation was set to 1 when the cell was mobile and co-localization was detected or when the cell was immobile and no co-localization was measured. In all the other cases, the correlation was set to -1 (Extended Data Fig. 7f). Then, a correlation score was attributed to each cell and each type of cluster (stationary or dynamic) by calculating the mean of all the values over time.

To further estimate the connection between cell movement and protein co-localization, a more quantitative analysis was performed. For each cell, speeds were first renormalized between 0 (no movement) and 1 (maximum velocity) and segmented into 15 equally spaced bins. Next, for each bin, we selected all the occurrences associated to a moving cell and verified whether static or dynamic co-localized clusters were detected. By repeating this analysis on all cells, we could therefore estimate for each bin the percentage of events where co-localization was detected. Finally, the percentage of co-localization was plotted separately for static and dynamic clusters as a function of the normalized cell speed.

Correlation between cell movement and cluster fluorescence intensity. To investigate the connection between cell movement and protein recruitment at focal adhesion complexes, we analysed how the intensity of AglZ and AglQ clusters changed with cell speed. Owing to cell-to-cell variability in the measured fluorescence signal, the intensity of static clusters was renormalized for each cell. For the cells displaying at least one dynamic cluster, the maximum intensity of dynamic clusters was calculated and used as an internal reference to normalize the intensity of all detected clusters within the cell. Behind this normalization procedure, we make the assumption that dynamic clusters have on average the same stoichiometry or composition from cell to cell. This assumption was verified since, after analysing more than 100 cells for AglZ and AglQ proteins, the intensity measured after normalization for the dynamic clusters showed a very small dispersion (mean intensity of 1 ± 0.15 for AglZ and AglQ) while the dispersion for the static clusters was about three times higher (mean intensity of 1.38 ± 0.5 for AglZ and 1.4 ± 0.43 for AglQ). Then, as already described for the correlation between cell movement and co-localization, we plotted the intensity of both static and dynamic clusters as a function of the normalized cell speed.

Note that this analysis could only be performed for AglZ and AglQ, as in this case the proportion of cells displaying both dynamic and static clusters was large enough to obtain a statistically relevant sample size (59% for AglZ and 64% for AglQ). For GltC and GltD, however, very few dynamic clusters were detected and these proportions dropped to 5% and 8%, respectively, which was insufficient for performing a similar analysis.

Analysis of TADA clusters by astigmatism. Detection of beads from the calibration were analysed using RapidSTORM³⁸. For each axial position (z) of the sample, the PSF of a single fluorescent bead was fitted with an elliptical Gaussian function and the x - y widths (w_x and w_y) were calculated in order to produce the calibration curves $w_x(z)$ and $w_y(z)$.

Only gliding cells displaying a bright and well-defined TADA cluster were selected for the analysis. First, the calibration curves were used to infer the axial positions of the clusters using RapidSTORM. Then, the cell outline was calculated on each image using the bright-field images. The cell outline was used to reconstruct the complete trajectory of the cell during the time-lapse and infer the lateral position of the TADA cluster. From the trajectory, we precisely determined for each image the total distance travelled by the cell since the beginning of the acquisition. Finally, the 3D position of the TADA clusters was plotted as a function of the distance travelled by the cell.

Principal component analysis and statistical tests. The PCA analysis was performed using Matlab 2014 and the Statistics and Machine Learning Toolbox. We started by defining a data matrix C_i for each mutant (WT, dA, dB, dC, ...) and reporter (AglZ, AglQ and GltI). For example, in the case of AglZ as a reporter, we defined one array C_{WT} for the wild-type strain and twelve others (C_{dA} , C_{dB} , C_{dC} , ...) for each mutant. Each row of an array C_i corresponded to a single observation of a cluster measured by fluorescence microscopy. The columns corresponded to the variables used to describe the properties of the clusters and of the associated strain: (i) longitudinal position of the cluster (0 at the leading pole, 1 at the lagging pole); (ii) lateral position (0 at the centre, 1 on the edge); (iii) total number of clusters detected simultaneously in the same cell; and (iv) proportion of cells containing at least one cluster for the selected strain and reporter.

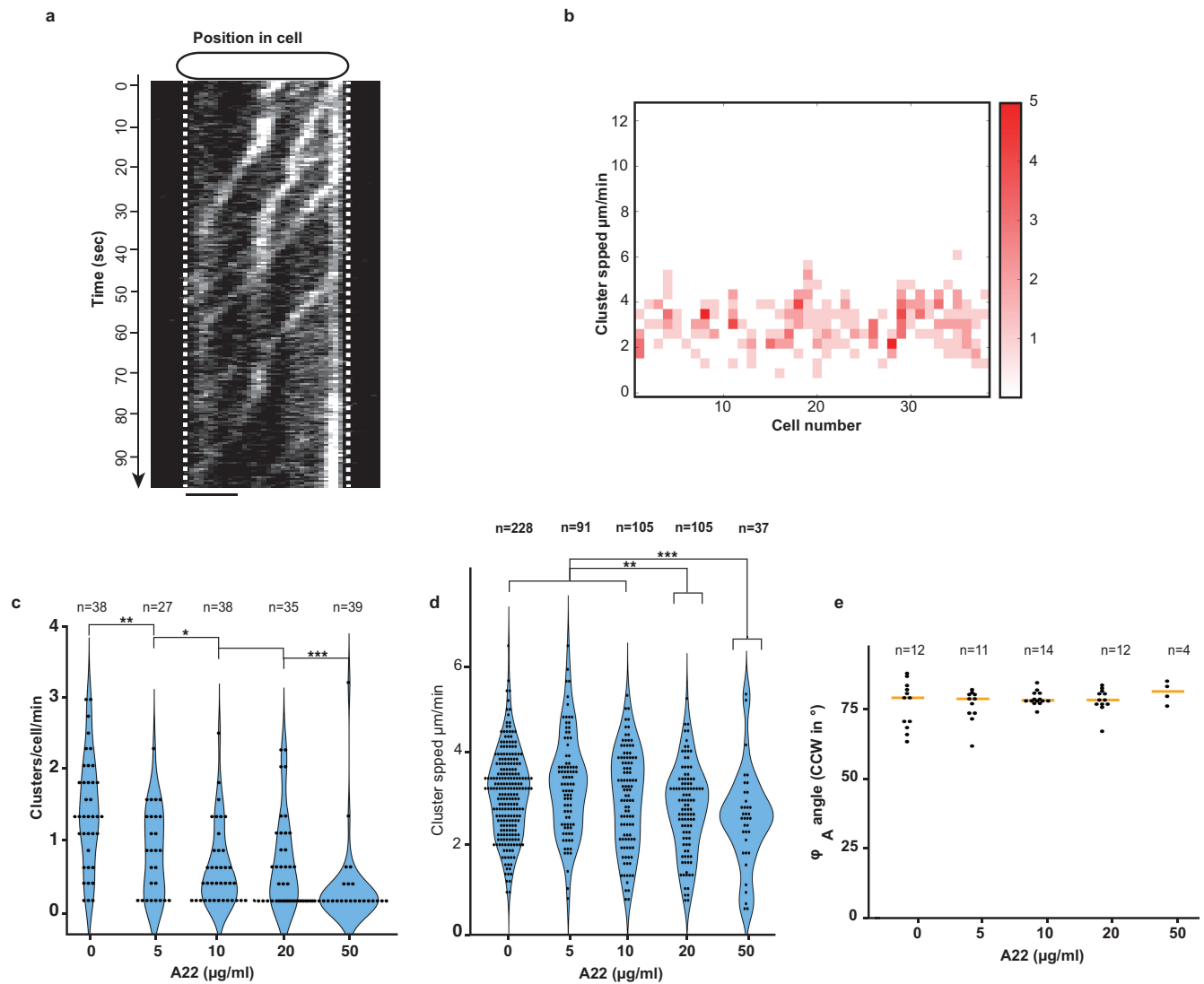
For each reporter, the PCA calculation was performed by combining all the data collected (wild-type and mutated strains). Therefore, a single set of principal components were defined and subsequently used to illustrate the differences between the wild-type and the mutated strains. The analysis was performed in three steps. First, all the data-matrices C_i were concatenated in a single matrix C_{all} and the values of each column were centred (mean was equal to 0) and standardized using the inverse variance. Then, the PCA analysis was performed and returned the coefficients of the principal components (PCs, Extended Data Fig. 4a–c) as well as the amount of variance accounted for by each of them (Extended Data Fig. 4d). Independently of the nature of the reporter, we observed that the three first components accounted for more than 87% of the total variance (>68% for the two first components). Therefore, the clusterization calculation was performed later using only the three first PCs.

In a second step, the values of the PCs were used to calculate the representations C_i^{PC} of each data-matrix C_i in the principal component space. Using a system of axes defined by the three first PCs, we could represent each array C_i^{PC} by a scatter plot in which each point represents the properties and localization of one single cluster in the PC space (Extended Data Fig. 4e). To improve the readability of the plots, we represented only the median position of each scatter plot (Extended Data Fig. 4e). We also used the standard deviation to illustrate the dispersion of the data along each PC. Finally, in order to make the comparison between wild-type and mutated strains easier, we arbitrarily placed the wild type at the origin.

Statistics and replicates. By default we used the Wilcoxon (two-sided) test to significantly separate the different samples. For each experiment, the number of times it was independently replicated in the laboratory (biological replicate) is indicated either in the figure legend or in the corresponding Methods section. All errors calculated on values were determined using the deviation standard formula.

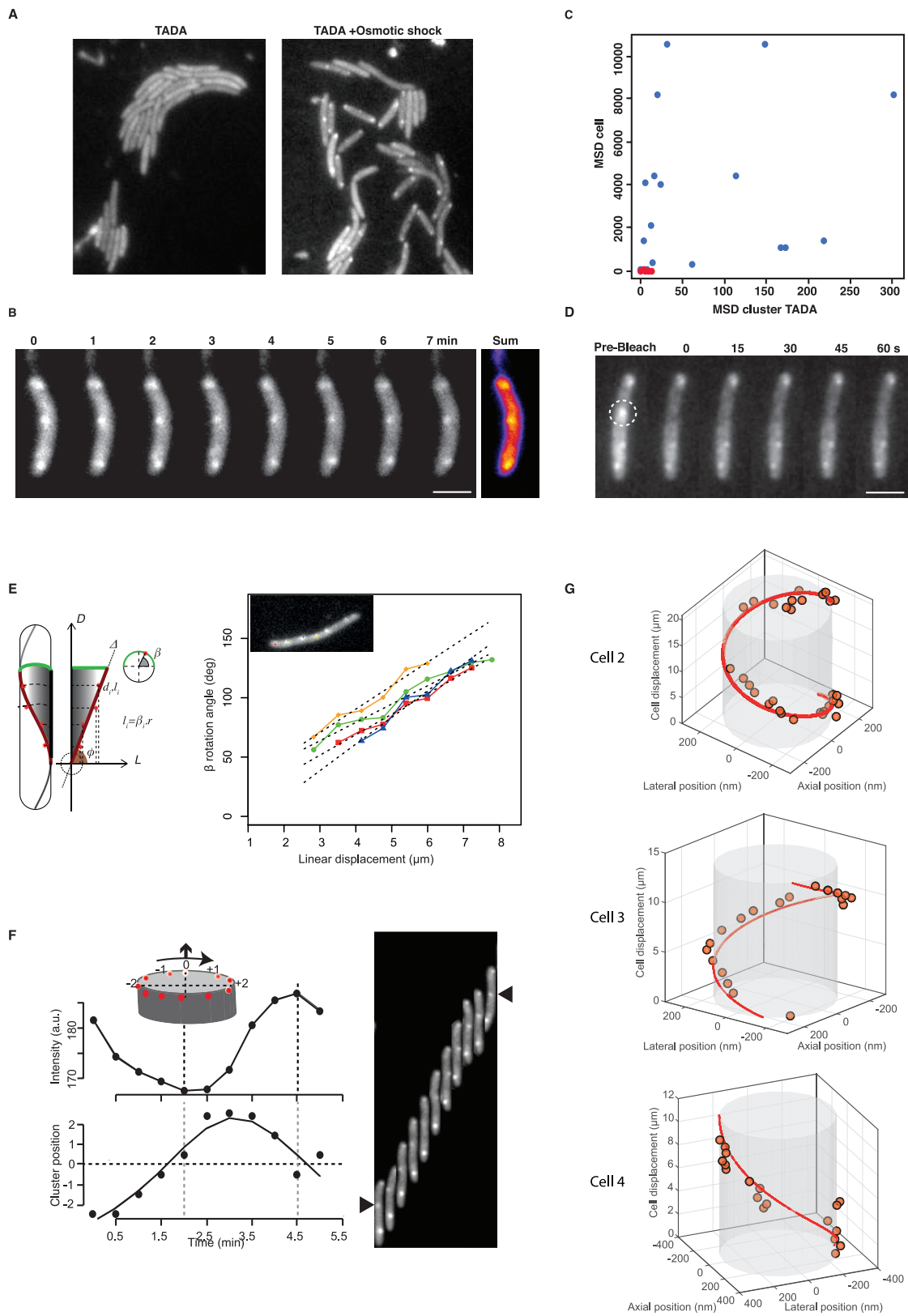
Data and code availability. Code used for two-colour TIRF and astigmatism experiments was written using Matlab 2015. Scripts are available upon request. The data that support the findings of this study are available from the corresponding author upon request.

- Bustamante, V. H., Martínez-Flores, I., Vlamakis, H. C. & Zusman, D. R. Analysis of the Frz signal transduction system of *Myxococcus xanthus* shows the importance of the conserved C-terminal region of the cytoplasmic chemoreceptor FrzCD in sensing signals. *Mol. Microbiol.* **53**, 1501–1513 (2004).
- Ducret, A., Théodoly, O. & Mignot, T. Single cell microfluidic studies of bacterial motility. *Methods Mol. Biol.* **966**, 97–107 (2013).
- Altschul, S. F. et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
- Johnson, L. S., Eddy, S. R. & Portugaly, E. Hidden Markov model speed heuristic and iterative HMM search procedure. *BMC Bioinformatics* **11**, 431 (2010).
- Söding, J., Biegert, A. & Lupas, A. N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* **33**, W244–W248 (2005).
- Holm, L., Kääriäinen, S., Rosenström, P. & Schenkel, A. Searching protein structure databases with DALI Lite v.3. *Bioinformatics* **24**, 2780–2781 (2008).
- Lassmann, T., Frings, O. & Sonnhammer, E. L. L. Kalign2: high-performance multiple alignment of protein and nucleotide sequences allowing external features. *Nucleic Acids Res.* **37**, 858–865 (2009).
- Pei, J., Sadreyev, R. & Grishin, N. V. PCMA: fast and accurate multiple sequence alignment based on profile consistency. *Bioinformatics* **19**, 427–428 (2003).
- Drozdetskiy, A., Cole, C., Procter, J. & Barton, G. J. JPred4: a protein secondary structure prediction server. *Nucleic Acids Res.* **43** (W1), W389–W394 (2015).
- Finn, R. D. et al. Pfam: the protein families database. *Nucleic Acids Res.* **42**, D222–D230 (2014).
- Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305**, 567–580 (2001).
- Käll, L., Krogh, A. & Sonnhammer, E. L. A combined transmembrane topology and signal peptide prediction method. *J. Mol. Biol.* **338**, 1027–1036 (2004).
- Marbouty, M. et al. Condensin- and replication-mediated bacterial chromosome folding and origin condensation revealed by Hi-C and super-resolution imaging. *Mol. Cell* **59**, 588–602 (2015).
- Ducret, A., Valignat, M. P., Mouhamar, F., Mignot, T. & Théodoly, O. Wet-surface-enhanced ellipsometric contrast microscopy identifies slime as a major adhesion factor during bacterial surface motility. *Proc. Natl Acad. Sci. USA* **109**, 10036–10041 (2012).
- Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
- Arigovindan, M. et al. High-resolution restoration of 3D structures from widefield images with extreme low signal-to-noise-ratio. *Proc. Natl Acad. Sci. USA* **110**, 17344–17349 (2013).
- Marques, O. *Practical Image and Video Processing using MATLAB* (John Wiley & Sons, 2011).
- Wolter, S. et al. rapidSTORM: accurate, fast open-source software for localization microscopy. *Nat. Methods* **9**, 1040–1041 (2012).



Extended Data Figure 1 | Dynamic AglZ-YFP clusters in non-motile cells. **a**, Dynamic AglZ-YFP clusters in a non-motile cell observed by TIRFM. Kymograph representation of cluster movement captured every 0.5 s. Note that the clusters form at the cell pole and move directionally towards the opposite cell pole where they are dispersed. Scale bar, 2 μm . **b**, Distribution of cluster speeds between and within cells. Note that clusters can move at different speeds in a cell and that the speed between

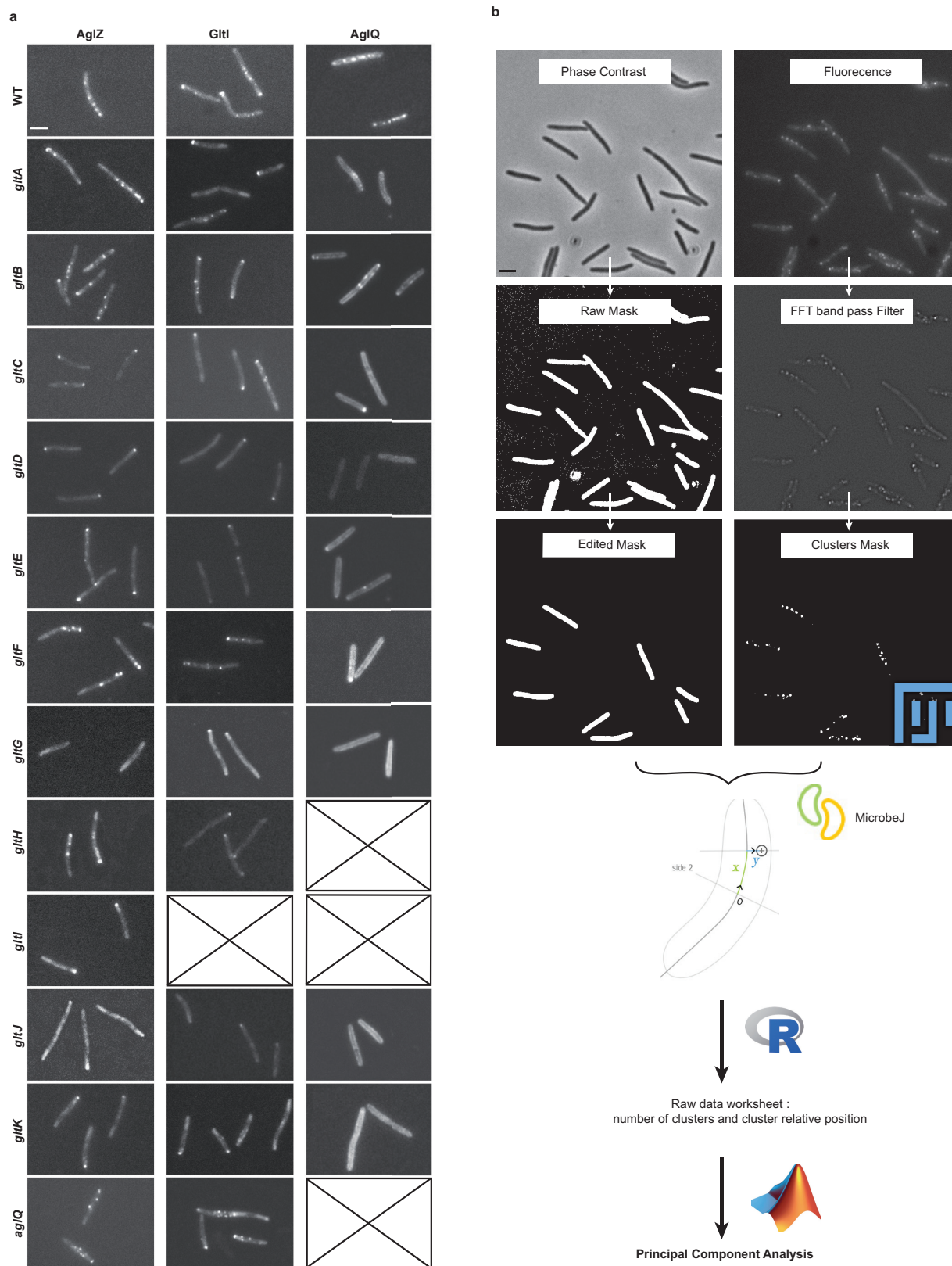
cells generally varies between 2 and 4 $\mu\text{m min}^{-1}$ (two technical replicates). **c**, Number of AglZ-YFP clusters per cell per minute in wild-type and A22-treated cells (two technical replicates). **d**, AglZ-YFP cluster speed in wild-type and A22-treated cells (two technical replicates). **e**, Trajectory angles in wild-type and A22-treated cells (two technical replicates). Wilcoxon tests; * $P < 0.1$; ** $P < 0.01$; *** $P < 0.001$ (c–e).



Extended Data Figure 2 | See next page for caption.

Extended Data Figure 2 | *Myxococcus* cells rotate along their long axes during motility. **a**, TADA-bright clusters form in *Myxococcus* cells subjected to a brief osmotic shock. TADA is incorporated in the *Myxococcus* cell envelope only when the cells are subjected to NaCl treatment (see Methods). **b**, **c**, TADA-bright clusters are not dynamic in non-motile cells. TADA-bright cluster movements are not detectable in non-motile cells (**c**, red dots, $n = 14$; two technical replicates) and only detectable in moving cells (**c**, blue dots, $n = 17$; two technical replicates). MSD, mean square displacement. Scale bar, $2\ \mu\text{m}$. **d**, TADA-bright clusters are inert. FRAP analysis reveals the absence of fluorescent molecule exchange in TADA-bright clusters ($n = 6$; one technical experiment). Scale bar, $2\ \mu\text{m}$. **e**, TADA-bright cluster rotation reflects rotation of the cell during movement. A cell on which four TADA-bright clusters were tracked is shown. The radial velocity of each cluster calculated by

projection of the 2D images on the model 3D cell cylinder (left, β angle) is plotted against the linear displacement of the cell. Each TADA cluster moved at the same radial speed and proportionally to the speed of the cell, indicating that TADA clusters are inert objects reporting on the rigid-body movement of the cell. **f**, Three dimensional trajectories of TADA-bright clusters reconstructed by astigmatism. In the absence of astigmatism, the size of TADA clusters corresponds to the diffraction limit of light, and they are circular (that is, the size of the PSF is the same in perpendicular directions), making the astigmatic analysis of axial position possible. **g**, TADA-bright clusters rotate clockwise. Cluster intensity fluctuations and positions relative to the cell axis are shown over time (left) in a representative cell (right). The black arrow points to the analysed cluster. The representative cell shown in the right panel was isolated from others in the field with a black mask ($n = 10$; three technical replicates).



Extended Data Figure 3 | Analysis of AglZ–YFP, GltI–YFP and AglQ–mCherry in *agl* and *glt* mutant backgrounds. **a**, Each fluorescent functional fusion gene was introduced in place of the wild-type gene in each genetic background shown. Typical examples are shown for each strain. Crossed boxes indicated genetic backgrounds that were not obtained for this study. Scale bar, 2 μ m. **b**, Cluster detection and analysis chart. Phase contrast and fluorescence images were processed so as to respectively extract cell masks of isolated cells (compared edited mask

to raw mask) and the position of fluorescence clusters following the application of a fluorescence bandpass filter. Note that the intensity of the fluorescence clusters was not exploited owing to lack of robustness and day-to-day fluctuations. The cluster coordinates were then defined relative to cell coordinates with the Microbe J plugin (<http://www.indiana.edu/~microbej/>) in Fiji, compiled in R sheets and further analysed by PCA using custom-written code in Matlab.

a

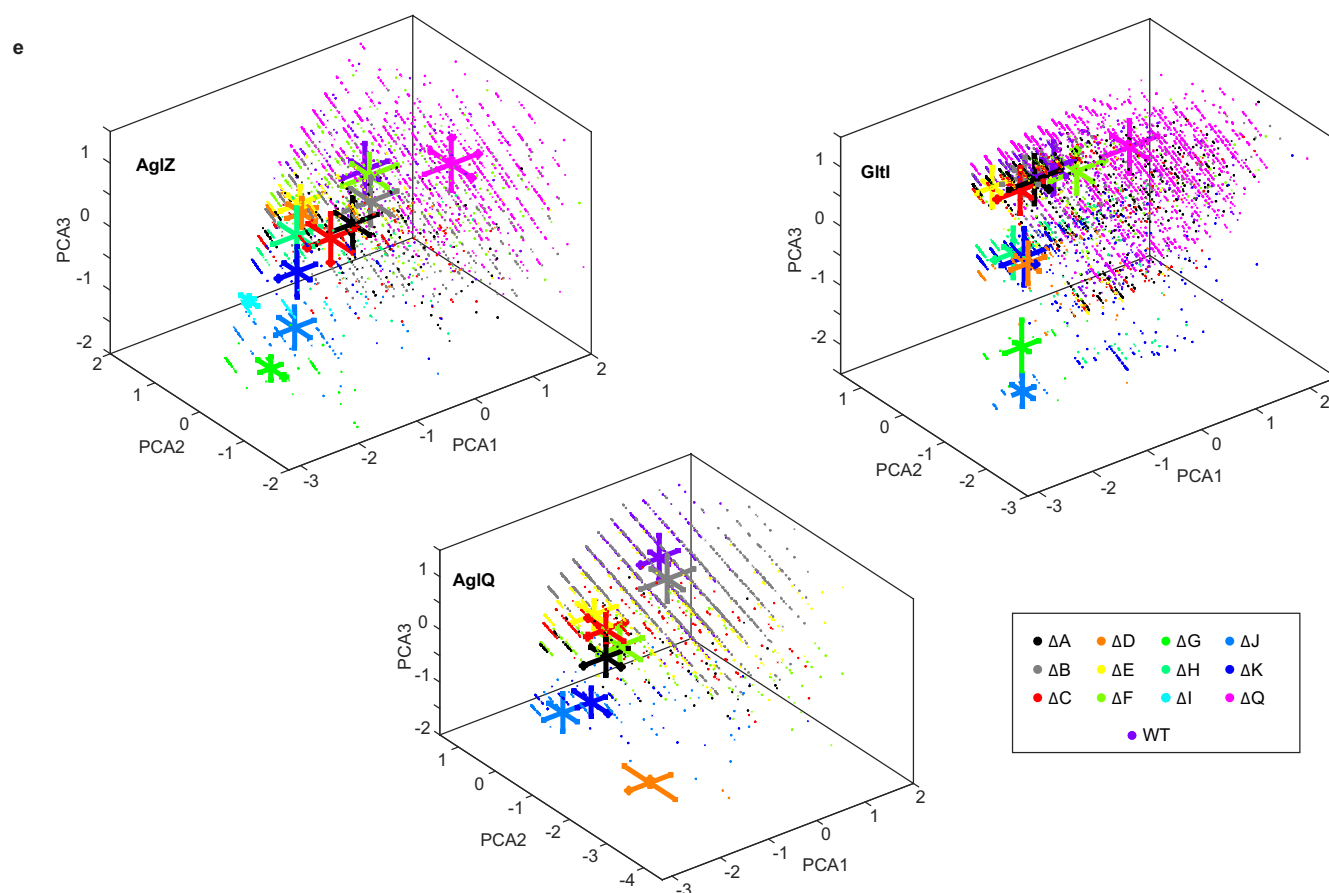
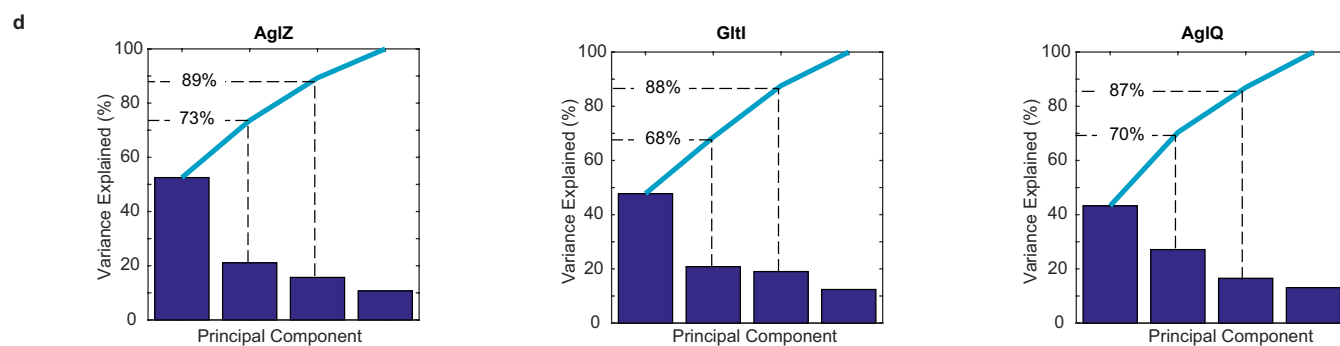
AgIZ	PC1	PC2	PC3	PC4
Longitudinal Position	0.5677	0.2464	-0.1929	0.7614
Axial Position	0.5054	-0.0425	0.8491	-0.1479
Number of cluster	0.3732	-0.8831	-0.2773	-0.0627
Proportion of cells	0.5320	0.3970	-0.4062	-0.6280

b

Giti	PC1	PC2	PC3	PC4
Longitudinal Position	0.5953	0.0891	-0.0150	0.7984
Axial Position	0.4817	-0.1351	0.8009	-0.3290
Number of cluster	0.4377	-0.7101	-0.4883	-0.2563
Proportion of cells	0.4712	0.6852	-0.3462	-0.4342

c

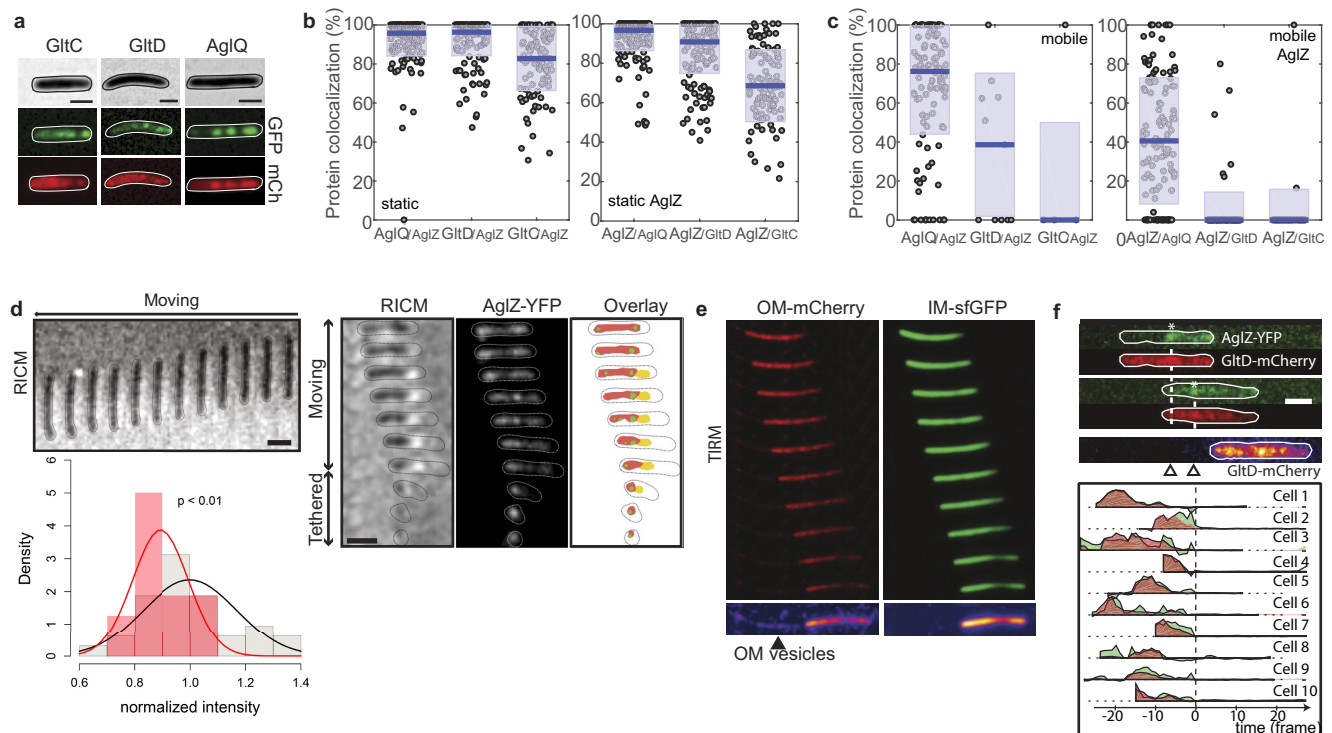
AgIQ	PC1	PC2	PC3	PC4
Longitudinal Position	0.6142	-0.0131	0.3934	0.6840
Axial Position	0.5559	0.1406	-0.8189	-0.0255
Number of cluster	0.2323	0.8417	0.3138	-0.3729
Proportion of cells	0.5098	-0.5211	0.2760	-0.6264



Extended Data Figure 4 | See next page for caption.

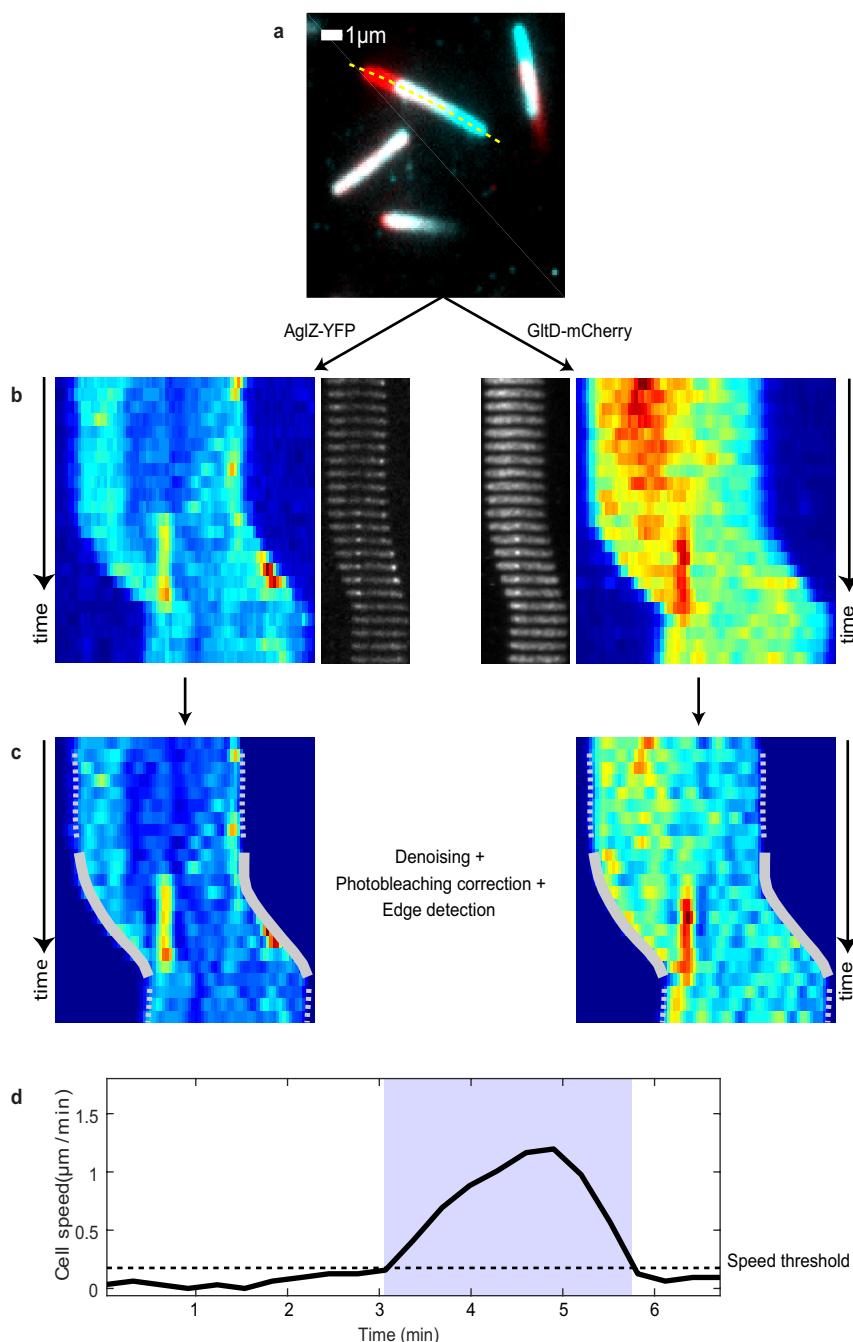
Extended Data Figure 4 | PCA of AglZ-YFP, GltI-YFP and AglQ-mCherry in *agl* and *glt* mutant backgrounds. **a–c,** Coefficients of the principal components (PCs) for AglZ-YFP, GltI-YFP and AglQ-mCherry. PCs are the eigenvectors of the correlation matrix calculated from the four parameters indicated in the first column of the tables. Together, PCs form an orthogonal basis in which the vectors are uncorrelated. PCs are sorted according to the amount of variability in the data they describe, PC1 having the largest effect (that is, variance) and PC4 the least. **d,** Scree plots displaying the variance associated to each PC. The bar plot represents

the variance associated with each PC for a given fusion (from left to right: AglZ-YFP, GltI-YFP and AglQ-mCherry). The cumulative variance is also plotted (light blue line). Note that PC1–2 describe on average 70% of the total variance and PC1–3 more than 87%. **e,** Projection of the data in the space defined by the three first PCs. For each mutant, data are represented by a scatter plot of a specific colour (see inset for colour code). For each direction and each mutant, the average and s.d. of the data are symbolized by a single bold line: the centre of the line represents the average and its length the s.d.



Extended Data Figure 5 | Motility is propelled by cyclic interactions between the inner membrane-localized motor and outer membrane-localized adhesins of the motility complex. **a**, Epifluorescence analysis of representative cells expressing AglZ-YFP-AglQ/GltD/GltC-mCherry. In each case, fluorescent functional fusions are expressed in place of the wild-type gene. Note that while AglZ-YFP and AglQ-mCherry clusters can be detected, GltC-mCherry and GltD-mCherry appear mostly diffuse around the cell envelope with these imaging conditions. Forty cells were imaged for AglZ-YFP-AglQ with two biological replicates; 232 cells were imaged for AglZ-YFP-GltD with three biological replicates; and 55 cells were imaged for AglZ-YFP-GltC with three biological replicates. Scale bar, 2 μm. **b**, Protein co-localization in static clusters. For each cell analysed, a percentage of co-localization is computed for proteins detected only in static clusters. Values range from 0 when no co-localization between the two proteins was detected in the cell to 100% when the two proteins were always detected together. The left panel shows the percentage of co-localization for AglQ, GltD or GltC with respect to AglZ. Single data are represented by a scatter plot (o), the median co-localization value is symbolized by a blue line and the s.d. by light grey boxes. On average, 96% of AglQ clusters colocalized with AglZ ($n = 153$ clusters; two biological replicates), 96% of GltD ($n = 120$; seven biological replicates) and 83% for GltC ($n = 100$; three biological replicates). Inversely, the right panel represents the percentage of co-localization of AglZ with respect to AglQ (97%, $n = 152$; two biological replicates), GltD (91%, $n = 120$; seven biological replicates) and GltC (69%, $n = 100$; three biological replicates), respectively. **c**, Protein co-localization in mobile clusters. Box-plots read as in **b** and describe the percentage of co-localization for proteins detected only in dynamic clusters. The left panel illustrates the co-localization of AglQ, GltD or GltC with respect to AglZ (76%, $n = 106$, two biological replicates; 39%, $n = 11$, seven biological replicates; and 0%, $n = 4$, three

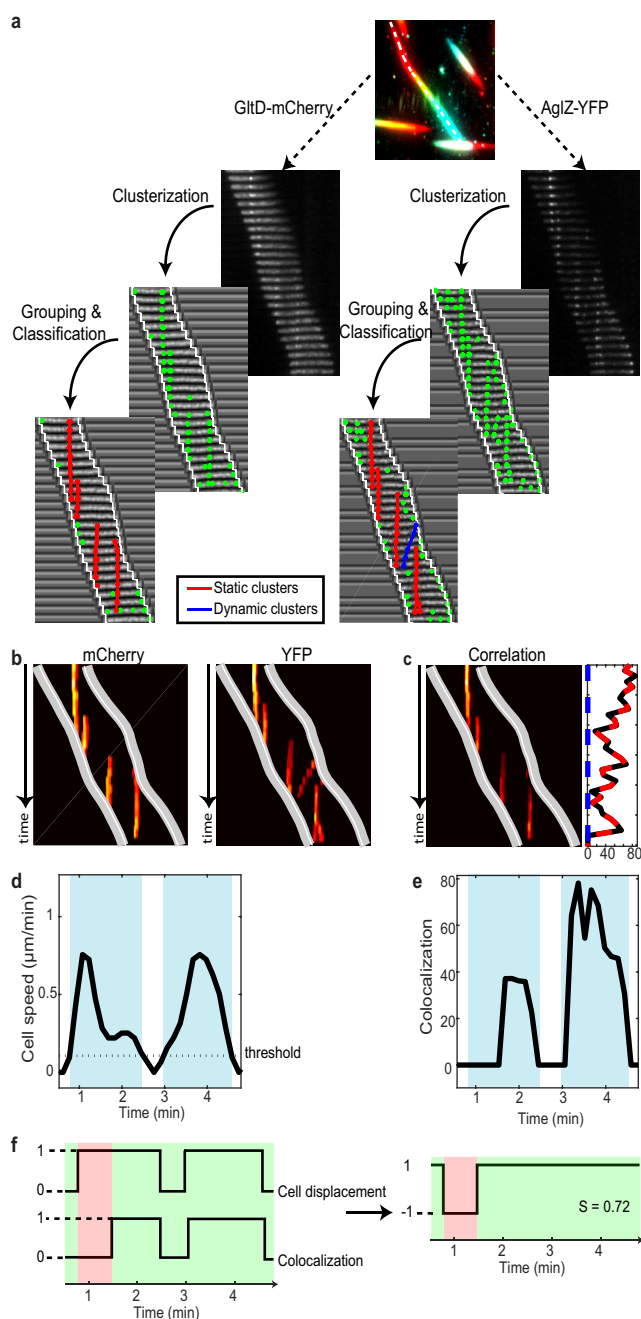
biological replicates, respectively). The right panel shows the percentage of co-localization of AglZ with respect to AglQ (41%, $n = 125$; two biological replicates), GltD (0%, $n = 72$; seven biological replicates) and GltC (0%, $n = 41$; three biological replicates). From the two panels, it appears that co-localization in dynamic clusters is essentially observed between AglZ and AglQ. For GltD and GltC, dynamic clusters are almost never detected ($n = 11$ for GltD, $n = 4$ for GltC) leading to a co-localization score close or equal to 0%. **d**, AglZ-YFP clusters localize within adhesive contact zones. Left, RISM of a representative gliding cell ($n = 10$, two biological replicates, 30-s time frames; scale bar, 2 μm) showing intimate connection with the chitosan-coated glass surface (dark zone). Right, adhesions and AglZ-YFP cluster localization in detaching cells by RISM and combined epifluorescence microscopy (time frames, 30 s; scale bar, 2 μm). The graph represents the distribution of RISM intensities at AglZ-YFP cluster positions (red line) compared to the average intensity along the whole cell body (black line). Data obtained for $n = 20$ cells; two biological replicates. **e**, Gliding *Myxococcus* cells deposit outer membrane vesicles in their wake. TIRFM images of a motile cell expressing both outer membrane sfGFP and inner membrane mCherry. Outer membrane vesicles are deposited, suggesting that the cell is firmly adhered to the underlying surface. Shown is a representative cell ($n = 60$; 12 technical replicates). **f**, GltD-mCherry, a periplasmic Glt protein, is not released by gliding cells at focal adhesion complexes. Top, TIRFM snapshots of a representative cell expressing both GltD-mCherry (red) and AglZ-YFP (green). The position of the GltD clusters on the surface coincides with position of focal adhesion complexes (white asterisk). Time frames, 15 s; scale bar, 2 μm. Bottom, variation of intensity for GltD-mCherry (red) and AglZ-YFP (green) as a function of time before (negative time) and after (positive time) the cell moved away from the focal adhesion position, shown for $n = 10$ cells (two biological replicates).



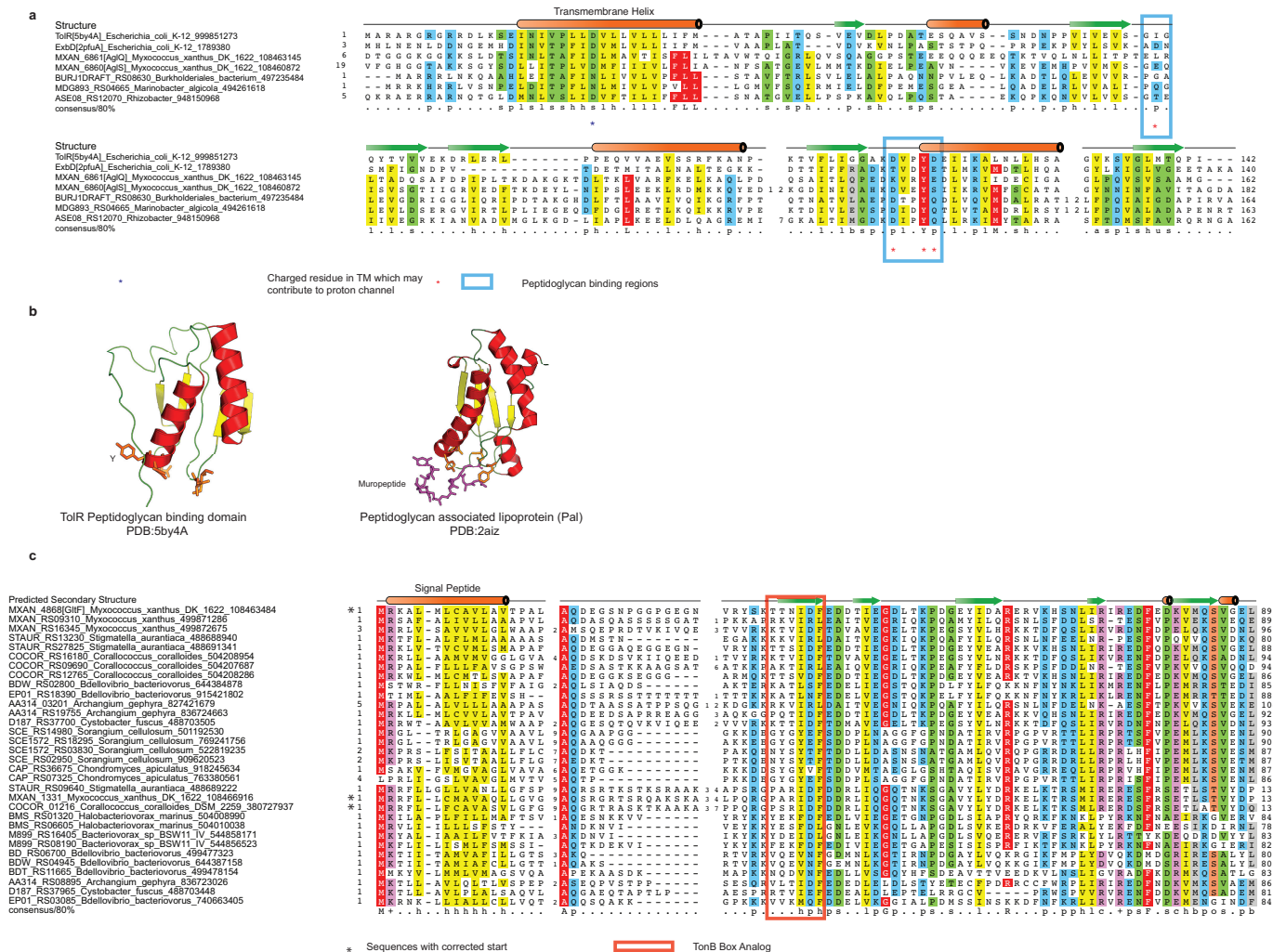
Extended Data Figure 6 | Image analysis for TIRFM experiments.

a, Temporal RGB image computed from fluorescence images (mCherry for this example). Images were summed together and colour coded from blue (first image) to red (last image). Immobile cells appear uniformly white whereas moving cells show coloured extremities. Cell trajectory is represented by a yellow dotted line. **b**, For the two imaging channels (YFP and mCherry), a kymograph and a montage were calculated. Kymograph reads from top (first image) to bottom (last image), each line representing the average fluorescent intensity computed along the cell trajectory. The montage shows for each acquisition an image of the cell after applying a

straightening algorithm. Clusters of proteins (AglZ-YFP or AglQ/GltD/GltI-mCherry) appear as bright spots at the centre of the cell. **c**, Kymographs after applying a denoising algorithm. The cell outline was depicted either by a dotted white line when the cell was immobile or by a continuous white line when the cell was gliding on the surface. **d**, Cell speed as a function of acquisition time. For each cell, a threshold was defined that depended on the SNR and the sample lateral stability during TIRFM acquisition. When the cell speed was below this threshold (horizontal dotted line), the cell was considered immobile. Phases associated with cell movement are highlighted in blue.



Extended Data Figure 7 | Co-localization estimation and correlation with cell movement. **a**, Co-localization between YFP-tagged and mCherry-tagged proteins was calculated in three steps. For each channel (YFP and mCherry), a threshold was applied to the montage in order to detect protein clusters (clusterization). When the same cluster was observed in successive images, its localizations were stitched together manually and finally classified as either static (red) or dynamic/mobile (blue). **b**, For each channel, a protein-detection map was computed from the kymograph and the positions of the detected clusters. The cell outline was depicted by two white lines. **c**, The co-localization map was obtained by multiplying the two protein-detection maps. In the inset, the cumulative intensities associated with static (red) and dynamic (blue) clusters are plotted as a function of time. **d**, Cell speed as a function of acquisition time. Blue boxes represent regions in which cell speed is higher than the threshold. **e**, Cumulative co-localization intensity (static and dynamic) as a function of acquisition time. Blue boxes represent regions with high cell speeds. **f**, Curves from **d** and **e** were binarized. For the cell speed, the value was set to '0' when the speed was below the threshold (immobile cell) and to '1' when it was above (gliding cell) (left). For the co-localization, the value was set to '1' when co-localization was detected, '0' otherwise (left). A correlation curve (right) was then computed by comparing the two curves. At each time-point, if the values of the two binarized curves were equal (1/1 or 0/0, green highlighted areas), the correlation was set to '1'. Otherwise, it was set to '-1' (red highlighted area). Finally, the correlation score was defined as the average of all the correlation values.

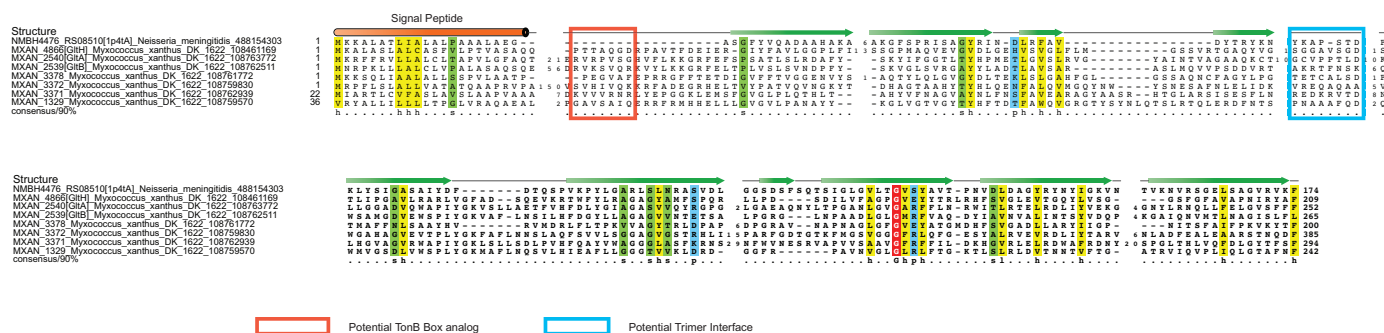


Extended Data Figure 8 | Bioinformatics analysis of AglQ/S and GltF.

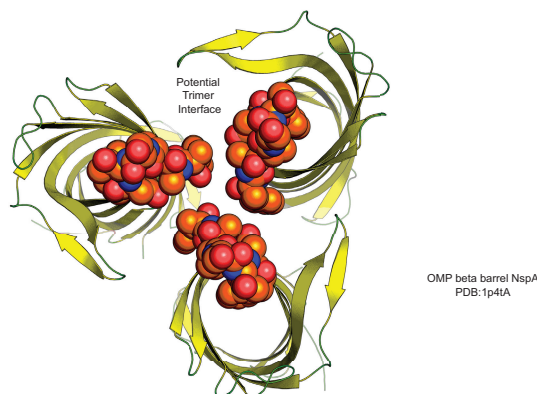
a, AglQ and AglS carry a potential peptidoglycan (PG)-binding site. Multiple alignment of AglQ, AglS and their paralogues with TolR(5by4A) and ExbD(2pfuA). The gene name, organism name and GI (NCBI) accession number are given. The structure is shown on top and the 80% consensus is shown below the alignment. **b**, Cartoon views of the structures are shown below with the residues known or predicted to bind peptidoglycan shown as sticks. **c**, The GltF family of proteins found in deltaproteobacteria. Multiple alignment of the GltF family is shown and

labelled using gene name, organism name and GI. The potential TonB box analogue is indicated. The TonB box can be most generally defined as an extended region that forms a β -strand-like structure that is not paired with other β -strands into a structural unit. The TonB box typically has two polar residues, T/S, and classically an acidic/amide residue. The GltF sequence profile analysis shows that GltF is related to the N-terminal region of certain OMP barrels that do contain a potential TonB-box like peptide.

a



b



Extended Data Figure 9 | Bioinformatics analysis of the GlT/A/B/H system. a, Multiple alignment of the β -barrel OMP proteins GlT/A, GlT/B and GlT/H and their paralogues with NspA (1p4tA). Note the presence of TonB box analogues in GlT/A and GlT/B but not GlT/H. **b,** Cartoon view of the inferred trimer of β -barrels based on the NspA structure with the residues predicted to form a trimeric interaction interface shown as spheres.

Doped polymer semiconductors with ultrahigh and ultralow work functions for ohmic contacts

Cindy G. Tang^{1,2*}, Mervin C. Y. Ang^{2,3*}, Kim-Kian Choo^{3*}, Venu Keerthi³, Jun-Kai Tan^{1,2}, Mazlan Nur Syafiqah³, Thomas Kugler⁴, Jeremy H. Burroughes⁴, Rui-Qi Peng^{1,2}, Lay-Lay Chua^{1,2,3} & Peter K. H. Ho^{1,2}

To make high-performance semiconductor devices, a good ohmic contact between the electrode and the semiconductor layer is required to inject the maximum current density across the contact. Achieving ohmic contacts requires electrodes with high and low work functions to inject holes and electrons respectively, where the work function is the minimum energy required to remove an electron from the Fermi level of the electrode to the vacuum level. However, it is challenging to produce electrically conducting films with sufficiently high or low work functions, especially for solution-processed semiconductor devices. Hole-doped polymer organic semiconductors are available in a limited work-function range^{1,2}, but hole-doped materials with ultrahigh work functions and, especially, electron-doped materials with low to ultralow work functions are not yet available. The key challenges are stabilizing the thin films against de-doping and suppressing dopant migration^{3,4}. Here we report a general strategy to overcome these limitations and achieve solution-processed doped films over a wide range of work functions (3.0–5.8 electronvolts), by charge-doping of conjugated polyelectrolytes^{5–7} and then internal ion-exchange to give self-compensated heavily doped polymers. Mobile carriers on the polymer backbone in these materials are compensated by covalently bonded counter-ions. Although our self-compensated doped polymers superficially resemble self-doped polymers^{8,9}, they are generated by separate charge-carrier doping and compensation steps, which enables the use of strong dopants to access extreme work functions. We demonstrate solution-processed ohmic contacts for high-performance organic light-emitting diodes, solar cells, photodiodes and transistors, including ohmic injection of both carrier types into polyfluorene—the benchmark wide-bandgap blue-light-emitting polymer organic semiconductor. We also show that metal electrodes can be transformed into highly efficient hole- and electron-injection contacts via the self-assembly of these doped polyelectrolytes. This consequently allows ambipolar field-effect transistors to be transformed into high-performance p- and n-channel transistors. Our strategy provides a method for producing ohmic contacts not only for organic semiconductors, but potentially for other advanced semiconductors as well, including perovskites, quantum dots, nanotubes and two-dimensional materials.

The usual way of making an ohmic contact involves using suitable metals or metal oxides as the electrode, which requires vacuum deposition and/or a high-temperature post-anneal step. Although charge-doped conducting polymers are promising alternatives, the availability of p- and n-doped polymer organic semiconductors is limited. Electro-oxidation and electro-reduction can generate a range of doped polymers, but are not generally practical for device applications. Surface dipolar layers, such as perfluorinated ionomers¹⁰ and polyethylenimine¹¹, have been reported to extend the work-function range of metal oxides, poly(3,4-ethylenedioxythiophene):poly(styrenesulfonic acid)

(PEDT:PSSH) films and sulfonated poly(thiophene-3-[2-(2-methoxyethoxy)ethoxy]-2,5-diyl):poly(4-hydroxystyrene) (S-P3MEET:PHOST) films, but they can introduce electrochemical effects and large electrical resistance, which degrade the quality of the contact^{12,13}.

Our self-compensated heavily doped polymer scheme uses a hole- or electron-doped polymer organic semiconductor core with charge compensation by covalently bonded counter-ions (Fig. 1a). Excess bound counter-ions can be used together with spectator ions to provide solution processability. The chemical structures and designations of these and other materials are shown in Extended Data Fig. 1. In contrast to self-doped polymers, these self-compensated doped polymers are produced by deliberate hole (or electron) doping with strong acceptors (or donors), and internal ion exchange to remove the inserted counter-ions as soluble salts (Extended Data Fig. 2). These processes can be performed in either the film state or the solution state; strong dopants can be used in a non-aqueous medium for quantitative doping. Near-unity doping efficiencies can be achieved because the doping is made irreversible by removal of the dopant product, which is not the case when the dopant is incorporated into the organic matrix¹⁴. Although self-compensation has previously been explored for charge balancing¹⁵, its potential to yield practical doped polymers with extreme work functions was not recognized. By contrast, polyelectrolytes have long been proposed as charge-transport layers and work-function modifiers, but their electrical resistance without deliberate doping is prohibitively high, and their carrier density, work function and charge-injection ability are limited^{5–7}. Therefore, ohmic injection of holes or electrons into semiconductors with high ionization potential (I_p) or low electron affinity (E_A), respectively, has not been possible.

Through the separate doping and compensation strategy, polymers with high I_p or low E_A can be heavily hole- or electron-doped, respectively, to reach a doping level well above 10 mol% of the repeat units and shift the Fermi level to the relevant band edges. This level of doping is different from that used for trap passivation, which requires only ultralow doping^{16–18}. Because the covalently bonded counter-ions prevent centre-of-mass diffusion, the doped carriers cannot diffuse as a neutral entity across the doped/undoped junction into the lower-energy semiconductor, despite favourable thermodynamics (Fig. 1b). This stabilizes the doping profile, that is, the doping density across the junction as a function of the depth or distance. To illustrate the stable doping profile achieved, the self-compensated hole-doped triarylamine-fluorene copolymer with tethered trifluoromethylsulfonylimide counter-ion (**P4**) was deposited on a substrate, and spin-cast with an overlayer film of an alkoxy-substituted poly(*p*-phenylenevinylene) (OC₁C₁₀-PPV), which has a lower I_p than **P4** (5.0 eV versus 5.8 eV). Optical spectroscopy demonstrates that the underlayer remains doped while the overlayer remains undoped (Fig. 1c). In contrast, if the counter-ion is unbonded, quantitative transfer of doping occurs to the OC₁C₁₀-PPV overlayer film (Extended Data Fig. 3a, b). Such a loss of

¹Department of Physics, National University of Singapore, Lower Kent Ridge Road, S117550 Singapore. ²Solar Energy Research Institute of Singapore (SERIS), National University of Singapore, 7 Engineering Drive 1, S117574 Singapore. ³Department of Chemistry, National University of Singapore, Lower Kent Ridge Road, S117552 Singapore. ⁴Cambridge Display Technology Ltd, Building 2020, Cambourne Business Park, Cambridge CB3 6DW, UK.

*These authors contributed equally to this work.

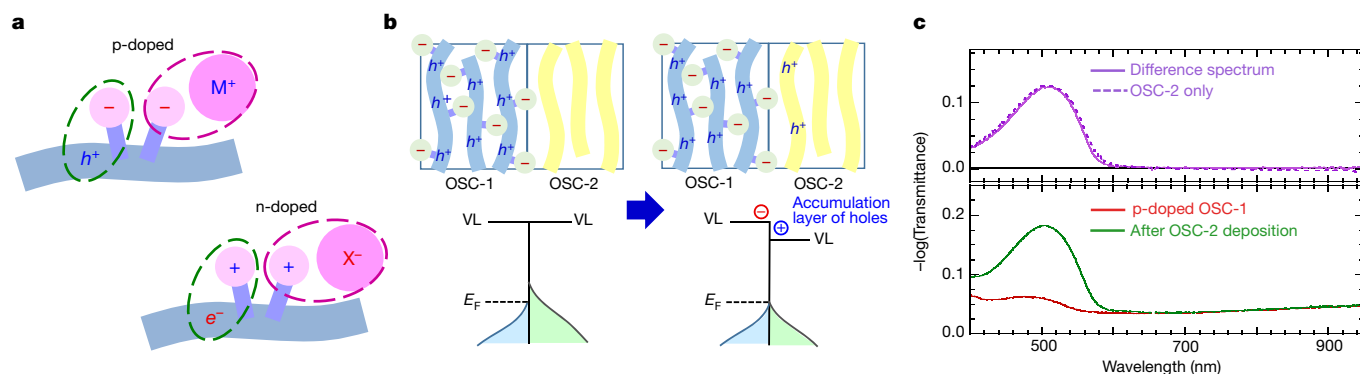


Figure 1 | Stabilization of the doping profile by counter-ion immobilization in self-compensated doped polymer organic semiconductors. **a**, Schematic of self-compensated p- (top) and n-doped (bottom) polymer organic semiconductors. The doped polymers contain a high density of mobile holes or electrons introduced into their π -conjugated cores. The green dashed ellipses highlight the tethered counter-ion required for self-compensation of the mobile carrier. The purple dashed ellipses highlight the tethered counter-ion and its spectator ion (M^+ or X^-) required for processability in solvents. **b**, **c**, Schematic (**b**) and experimental data (**c**) for stabilization against doping transfer from a self-compensated hole-doped organic semiconductor layer (OSC-1) with higher energy for the charge carriers to an adjacent organic semiconductor (OSC-2) with lower energy, during solution processing. Fermi-level

(E_F) alignment results in an offset of the vacuum level (VL) due to the transfer of carriers at the interface. In the experiment, OSC-2 is poly(2-methoxy-5-(3',7'-dimethyloctyloxy)-1,4-phenylenevinylene) (OC_1C_{10} -PPV; $I_p \approx 5.0$ eV), spin-cast over a triarylamine-fluorene copolymer film with tethered trifluoromethylsulfonylimide counter-ion (**P4**; OSC-1) ($I_p \approx 5.8$ eV) that was p-doped to the self-compensated state (doping level of $x \approx 0.8$). These films were probed by optical transmission spectroscopy. The difference spectrum (top panel, purple solid line) is obtained from the sample spectrum after depositing OSC-2 (bottom panel, green line) minus the spectrum before deposition (bottom panel, red line). The spectrum of OSC-2 only on glass (top panel, purple dashed line) is shown for comparison. No doping transfer to the OSC-2 overlayer has occurred.

control of the doping profile occurs whenever mobile counter-ions are present, even as impurities (Extended Data Fig. 3c, d). Only when all the counter-ions are immobilized will carrier diffusion occur solely across the heterojunction to produce the desired charge accumulation layer¹⁹.

Details of synthesis, doping and self-compensation of the exemplified materials are given in Supplementary Methods. The doping level of these materials was ascertained by optical spectroscopy and/or X-ray photoemission spectroscopy (XPS); see Supplementary Discussion 2.1 and 2.2. Self-compensation was verified by XPS at the film surface and by Fourier-transform infrared (FTIR) spectroscopy for the film bulk. The basic characterizations of selected materials are reported in Supplementary Discussion 2.2 and Extended Data Fig. 4. The vacuum work function was measured in the standard way by ultraviolet photoemission spectroscopy (UPS) from the Fermi level of the film.

The work function has an electrochemical and a surface-dipole contribution, represented by the sum $\mu + \chi$, where μ is the electrochemical potential of carriers at the Fermi level, measured from the Galvani potential, and χ is the surface-dipole potential²⁰. For doped organic semiconductors, the 'intrinsic' component of the work function, which is dominated by μ , is governed by the electronic structure of the organic semiconductor core; see ref. 21 for examples of hole-doped polymers and Extended Data Fig. 4f, j, k for examples of electron-doped polymers. However, Coulomb interactions with counter-ions and any spectator ions that are present also play an important part, owing to poor screening in these materials^{21,22}. On the other hand, the 'extrinsic' component of the work function, given by χ , can be manipulated through surface ionic layering and dipole effects.

As an illustration of the effects of the π -conjugated core, self-compensated hole-doped triarylamine-fluorene copolymers can provide ultrahigh work functions up to 5.8 eV. Self-compensated electron-doped naphthalene-bis(carboxyimide)-thiophene copolymers can provide low work functions down to 3.9 eV. These were the first examples of air-stable, heavily n-doped polymers that can be processed from solution. The use of fluorene copolymer cores can extend the work-function range to 3.0 eV under stringent conditions.

Doped organic semiconductor films with such extreme work functions are vulnerable to de-doping by impurities such as H_2O and O_2 , which are present even in glove boxes. If the Fermi level of the doped film is higher or lower than the redox potential of the O_2/H_2O couple,

electron transfer can occur to de-dope the film. The standard electrode potential for this couple in neutral H_2O corresponds to a work function of 5.3 eV. However, the actual conditions during film formation can be very different from the standard conditions used for these potentials. Also, the actual mechanism may involve considerable overpotentials, as in one-electron oxygen reduction and water oxidation reactions (Extended Data Fig. 5). As a consequence, the regime above 5.3 eV—which is demarcated by oxidation of neutral H_2O under standard conditions and which we denote 'ultrahigh work function'—and the regime below 4.1 eV—which is demarcated by standard reduction of O_2 to $O_2^{\cdot-}$ in water and which we denote 'ultralow work function'—may exhibit sufficient kinetic stability for device processing before encapsulation.

The stringent removal of H_2O is important for the stability of films with both ultrahigh and ultralow work functions. H_2O promotes the oxygen reduction cascade and autoxidation reactions at much less negative potentials, which closes the kinetic stability window. Therefore, these materials need to be non-hygroscopic. This condition can be met practically by selecting minimally hygroscopic counter-ions and spectator ions (Supplementary Information 2.3). For example, the trifluoromethylsulfonylimide counter-ion imparts considerably better ambient stability to the ultrahigh-work-function hole-doped *p*-trifluoromethyl-substituted triarylamine-fluorene copolymer than does the more hygroscopic sulfonate ion (Extended Data Fig. 6a, b). Furthermore, covalent bonding of the counter-ions confers stability against de-doping. For example, the ω -trimethylammonioalkyl counter-ion imparts considerably better ambient stability to the ultralow-work-function electron-doped naphthalenebis(carboxyimide)-thiophene copolymer than does unbonded Na^+ or cobaltocenium (Extended Data Fig. 6c, d), providing unprecedented, albeit limited, air stability to the self-compensated electron-doped naphthalene-bis(carboxyimide)-thienothiophene copolymer with tethered trimethylammonium counter-ion (**N1**; which has a vacuum work function of 4.0 eV). This ability of the tethered counter-ions to improve polymer stability is attributed to suppression of the ionic conduction mechanism that is needed to compensate electronic de-doping of the polymer^{2,23}.

These design considerations enable a wide variety of self-compensated heavily doped polymer organic semiconductors to be made to exhibit ultrahigh and ultralow work functions. With these materials, ohmic hole and electron contacts can be achieved even for the most

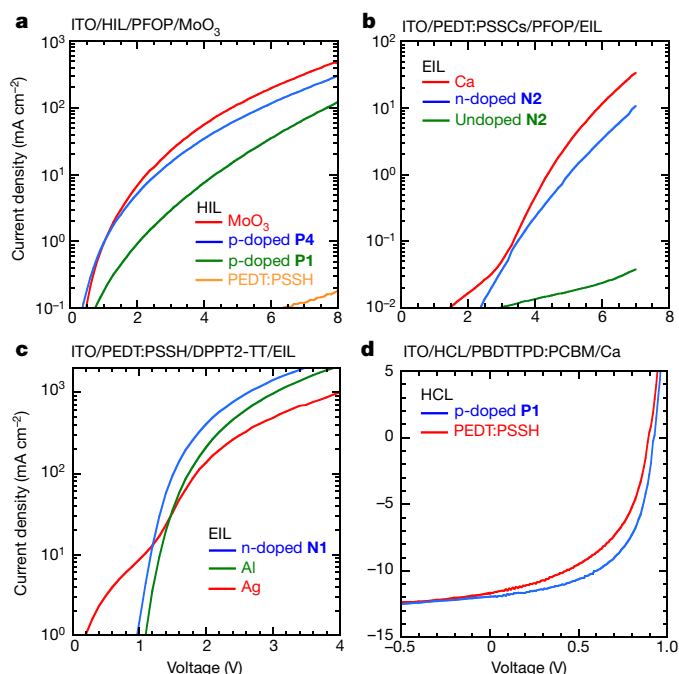


Figure 2 | Ohmic charge-injection and charge-collection characteristics. **a**, Current-density-voltage (J - V) characteristics of hole-only diodes (ITO/HIL/PFOP/MoO₃) with a 170-nm-thick PFOP layer. 30-nm-thick p-doped **P4** (blue line), p-doped **P1** (green line) or PEDT:PSSH (orange line), spin-cast onto ITO before the deposition of PFOP, was used as the HIL in these test diodes. 10-nm-thick MoO₃ capped with Ag was evaporated onto the top of the PFOP layer to give the control ohmic hole contact. The J - V characteristic obtained when biased for hole injection from MoO₃ is shown for comparison (red line). **b**, J - V characteristics of electron-only diodes (ITO/PEDT:PSSCs/PFOP/EIL) with a 120-nm-thick PFOP layer. 50-nm-thick p-doped PEDT:PSSCs, spin-cast onto ITO before the deposition of PFOP, was used as the hole-blocking contact. 20-nm-thick **N2**, either undoped (green line) or n-doped by contact with sodium naphthalenide and capped with evaporated Ag (blue line), was used as the EIL, which was spin-cast onto the top of the PFOP layer. 30-nm-thick Ca, capped with Al, was evaporated onto the top of the PFOP layer to give the control ohmic electron contact (red line). **c**, J - V characteristics of bipolar diodes (ITO/PEDT:PSSH/DPPT2-TT/EIL) with a 110-nm-thick DPPT2-TT layer. 50-nm-thick p-doped PEDT:PSSH, spin-cast onto ITO before the deposition of DPPT2-TT, was used as the ohmic hole contact. 5-nm-thick n-doped **N1** capped with Ag, spin-cast onto the top of the DPPT2-TT layer, was used as the EIL in the test diode (blue line). 100-nm-thick Ag (red line) and Al (green line), evaporated onto the top of the DPPT2-TT layer, were used as controls. **d**, J - V characteristics of solar cells (ITO/HCL/PBDTTPD:PCBM/Ca) with a 90-nm-thick PBDTTPD:PCBM (at a weight ratio of 1:1.5) photoactive layer, under 100 mW cm⁻² illumination (spectral-mismatch corrected). 40-nm-thick p-doped PEDT:PSSH (red line) or 30-nm-thick p-doped **P1** diluted in PHOST (at a weight ratio of 1:3) (blue line), spin-cast onto ITO before the deposition of the photoactive layer, was used as the HCL. 30-nm-thick Ca, evaporated onto the top of the photoactive layer, was used as the ECL.

demanding organic semiconductor benchmark—the wide-bandgap poly(9,9-bis(4-octylphenyl)fluorene-2,7-diyl) (PFOP). To demonstrate the creation of these contacts, we use unipolar single-carrier device protocols, thus avoiding any ambiguity related to the confinement of opposite carriers. For example, p-doped **P4** and MoO₃ are equally as effective at providing ohmic hole injection into PFOP ($I_p \approx 5.8$ eV) when used as the hole-injection layer (HIL) in the hole-only device ITO/HIL/PFOP/MoO₃/Ag (where ITO is indium tin oxide). An evaporated 10-nm-thick MoO₃ film provides an authentic top ohmic hole contact in these devices as reference. For device fabrication details, see Supplementary Methods. The current density is essentially the same whether biased for hole injection from **P4** or MoO₃ (Fig. 2a). The current-density-voltage (J - V) characteristics exhibit the ideal form

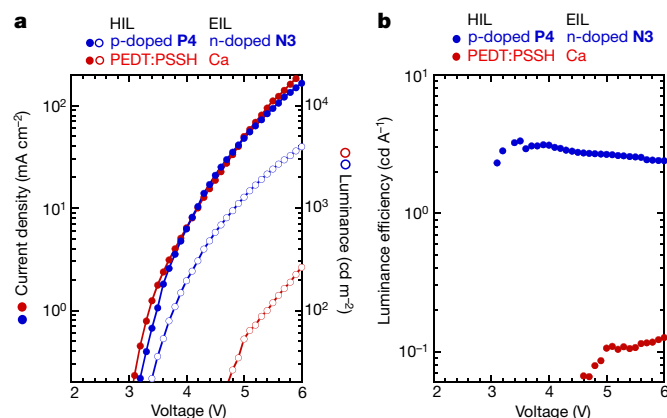


Figure 3 | Blue-light-emitting diode with ultrahigh- and ultralow-work-function charge injection layers. **a**, Current-density-voltage (J - V ; filled symbols, left axis) and luminance-voltage (L - V ; open symbols, right axis) characteristics for a 90-nm-thick PFOP light-emitting diode (ITO/HIL/PFOP/EIL). 30-nm-thick p-doped **P4** and 20-nm-thick n-doped **N3**, spin-cast before and after the deposition of the PFOP layer, were used as the HIL and EIL, respectively, in the test diode (blue symbols). A diode with 40-nm-thick p-doped PEDT:PSSH, spin-cast before the deposition of the PFOP layer, and 30-nm-thick Ca, evaporated onto the PFOP layer, was used as control (red symbols). All diodes were capped with Al. **b**, Luminance efficiency versus voltage for the diodes.

($J \propto V^2$) for space-charge-limited current conduction. Whereas MoO₃ has to be deposited in high vacuum, **P4** can be deposited from solution, before or after the semiconductor layer as desired and even under ambient conditions, which increases the flexibility of device design. When **P4** (work function of 5.8 eV) is substituted by the self-compensated hole-doped triarylamine-fluorene copolymer with tethered sulfonate counter-ion (**P1**) or by PEDT:PSSH (work functions of 5.4 eV and 5.1 eV, respectively), the hole-injection efficiency is greatly reduced. The effective work function evaluated *in situ* from built-in potential (V_{bi}) measurements using Stark-effect electro-absorption spectroscopy is broadly consistent with the vacuum work function (Supplementary Discussion 2.4 and Extended Data Fig. 7).

In parallel experiments, a self-compensated electron-doped benzothiazole-fluorene copolymer with tethered trimethylammonium counter-ion (**N2**; vacuum work function of 3.9 eV) is nearly as effective as Ca at providing ohmic electron injection into PFOP (estimated $E_A \approx 2.7$ eV) when used as the electron-injection layer (EIL) in the electron-only device ITO/PEDT:PSSCs/PFOP/EIL/Ag (where PEDT:PSSCs is poly(3,4-ethylenedioxythiophene):poly(styrenesulfonate cesium)). The PEDT:PSSCs layer (vacuum work function of 4.5 eV) is selected to be hole-blocking to ensure a unipolar device. An **N2** layer was spin-cast onto the PFOP as a putative EIL. It was left undoped, or n-doped with a dilute solution of sodium naphthalenide, and capped with evaporated Ag. Control devices were fabricated by evaporating Ca on PFOP. The J - V characteristics show that n-doped **N2** and Ca, but not undoped **N2**, are able to inject electrons into the semiconductor (Fig. 2b). The J - V characteristics are consistent with trap-limited space-charge-limited current; that is, $J = (V - V_0)^n$ where $V_0 \approx V_{bi}$ is the bare potential and $n = 4.8$ (ref. 24).

n-doped **N1** can be used with p-doped PEDT:PSSH to provide ohmic electron and hole contacts to low-bandgap organic semiconductors under ambient processing conditions. In combination with PEDT:PSSH, **N1** is equivalent to or better than Al as an EIL on poly(2,5-bis(alkyl)-1,4-dioxypyrrolo[3,4-c]pyrrole-3,6-diyl-thiophene-2,5-diyl-thieno[2,3-b]thiophene-2,5-diyl-thiophene-2,5-diyl) (DPPT2-TT) in the bipolar diode structure ITO/PEDT:PSSH/DPPT2-TT/EIL/Ag (Fig. 2c). DPPT2-TT is an important prototype of the low-bandgap DPP polymers ($I_p \approx 5.15$ eV; estimated $E_A \approx 3.5$ eV)²⁵. n-doped **N1** can also be an efficient bottom electron-collection layer (ECL) for poly(3-hexylthiophene):phenyl-C₆₁-butyrate methyl ester

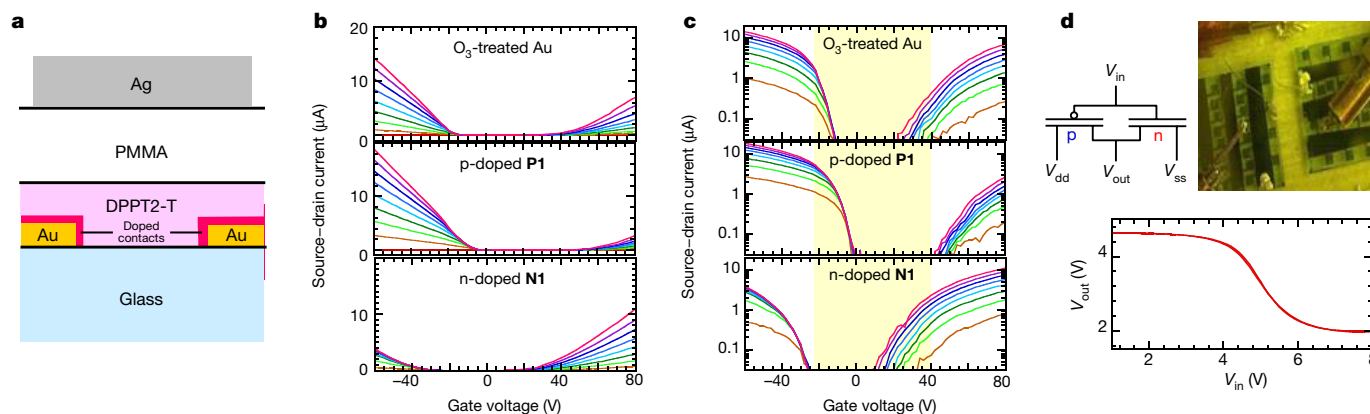


Figure 4 | Facile electrode differentiation by self-aligned assembly of ultrahigh- or ultralow-work-function interlayers. **a**, Schematic of the FET structure. Grey, Ag gate electrode; white, 450-nm-thick PMMA dielectric; pink, 50-nm-thick DPPT2-T organic semiconductor; magenta (in test devices), doped contacts imposed by a self-assembled ultrahigh- or ultralow-work-function interlayer; yellow, Au source/drain electrodes; blue, glass. **b**, **c**, Transfer characteristics on linear (**b**) and semi-logarithmic (**c**) scales. A monolayer of p-doped **P1** (middle) or n-doped **N1** (bottom) was self-assembled onto the Au electrode array from dilute solution.

(P3HT:PCBM) solar cells in the inverted architecture with PEDT:PSSH as top hole-collection layer (HCL) (Extended Data Fig. 8). This opens up the possibility of fabricating inverted organic and other solar cells on flexible metal foils.

p-doped **P1** outperforms the usual PEDT:PSSH as the HCL in model poly(5,6-dihydro-5-alkyl-4,6-dioxo-4*H*-thieno[3,4-*c*]pyrrole-1,3-diyl-4,8-bis(alkoxy)-benzo[1,2-*b*:4,5-*b'*]dithiophene-2,6-diyl):phenyl-C₆₁-butyrate methyl ester (PBDTPD:PCBM) organic solar cells of the form ITO/HCL/PBDTPD:PCBM/Ca/Al, with PBDTPD:PCBM at a weight ratio of 1:1.5. PBDTPD is an example of a donor polymer with a higher I_p and, hence, better ambient stability than regioregular P3HT²⁶. The cells with **P1** outperform those with PEDT:PSSH (Fig. 2d). Higher fill factor (0.60 versus 0.50; standard error of ± 0.02 , $n = 8$), open-circuit voltage (0.93 V versus 0.90 V) and power conversion efficiency (6.8% versus 5.2%) were obtained without any particular morphology optimization of the photoactive layer. The improvements in open-circuit voltage and, in particular, fill factor are characteristic signatures of the reduced contact resistance in highly ohmic contacts (J.-K.T., R.-Q.P. and P.K.H.H., unpublished results). These results demonstrate that, to achieve the required ohmic contacts and best performances from state-of-the-art photoactive layers, HCL materials with ultrahigh work function need to be used instead of PEDT:PSSH.

In a final experiment on diodes, we demonstrate solution-processed ohmic contacts to a model blue-light-emitting polymer diode. The device protocol is ITO/HIL/PFOP/EIL/Al. A 90-nm-thick PFOP layer was used as the model polymer organic semiconductor, 30-nm-thick p-doped **P4** as the HIL and 20-nm-thick self-compensated electron-doped fluorene polymer with tethered trimethylammonium counterion (**N3**; vacuum work function of 3.0 eV) as the EIL, capped with Al. This test diode gives a larger (by about 0.3 V) V_{bi} and a higher current density when plotted against $V - V_{bi}$, compared to the control device with PEDT:PSSH as the hole contact (which is blocking) and evaporated Ca as the EIL (Fig. 3a). The test diode exhibits more balanced electron-hole injection, which produces a 1.5-order-of-magnitude improvement in electroluminescence efficiency ($2\text{--}3\text{ cd A}^{-1}$) over the control device (Fig. 3b). This, together with the demonstration of efficient DPPT2-TT bipolar diode, demonstrates that fully solution-processed p-i-n structures can be produced. The deposition of the electron-doped polymer overlayer is compatible with the hole-doped polymer underlayer. Also, the doped layers do not degrade the electroluminescence properties of the sandwiched polymer organic semiconductor layer. These features are possible only because of the counter-ion

Channel length, 100 μm ; channel width, 3 mm. The drain voltage V_d was stepped from 0 V (red) to -20 V (magenta) in steps of -2.5 V. An O_3 -treated Au electrode array (top) was used as the control. The yellow shading in **c** marks the subthreshold region of the control device to emphasize the threshold shifts in the test devices. **d**, Schematic (top left), photograph (top right) and characteristics (bottom) of a complementary metal-insulator-semiconductor inverter. Common $V_{ss} = 0.0$ V, supply voltage $V_{dd} = 5.0$ V; V_{in} , input voltage; V_{out} , output voltage.

immobilization that prevents doping-profile migration and mutual annihilation of the opposite carriers.

Next we demonstrate the ability of self-assembled ultrahigh- and ultralow-work-function polymer interlayers to differentiate regular metal electrodes, such as Au and Ag, into excellent hole and electron contacts, respectively, in field-effect transistors (FETs). The top-gate bottom-contact FETs tested comprise Au source-drain electrodes fabricated on a glass substrate, poly(2,5-bis(alkyl)-1,4-dioxopyrrolo[3,4-*c*]pyrrole-3,6-diyl-terthiophene-2,5''-diyl) (DPPT2-T) semiconductor, poly(methyl methacrylate) gate dielectric and Ag gate electrode (Fig. 4a). Before deposition of the semiconductor, the Au electrodes were optionally contacted with a solution of p-doped **P1** or n-doped **N1** to self-assemble a doped-polymer monolayer that is a few nanometres thick. This assembly occurs over only the patterned metal electrodes, but not the channel, owing to image-charge electrostatic attraction.

The control FET with bare Au source-drain electrodes (ozone-cleaned) exhibits ambipolar behaviour with field-effect mobilities of $1.0\text{ cm}^2\text{ V}^{-1}\text{ s}^{-1}$ and $0.5\text{ cm}^2\text{ V}^{-1}\text{ s}^{-1}$ for holes and electrons, respectively. Insertion of **P1** suppresses n-channel behaviour through a 10-V upshift of the gate-source V_{gs} thresholds, whereas insertion of **N1** suppresses p-channel behaviour by a 15-V downshift of the V_{gs} thresholds (Fig. 4b, c). Such large shifts in both directions cannot be obtained with the usual thiol self-assembled monolayer treatment of Au²⁷. Although FETs are less sensitive than diodes to their contacts, owing to carrier accumulation in their channels²⁸, switching the work function across a 1.5-eV-wide range here is sufficient to convert the electrode from a good hole contact to a good electron contact, thereby differentiating ambipolar FETs into p-FETs and n-FETs. This switching is also observed for Ag electrodes, despite their intrinsically better electron-injection properties (Extended Data Fig. 9). It is thus a general approach to contact modification. This ability to manipulate the work function of metal electrodes will help to realize optimal contacts also to two-dimensional semiconductors²⁹.

The differentiated p-FET and n-FET can be combined to give an inverter—the most basic element of a complementary metal-insulator-semiconductor circuit (Fig. 4d). The device shows clean inversion operation, no hysteresis, and a midpoint voltage of 5 V that is slightly displaced by asymmetry effects. Other possible applications include vertical transistors and light-emitting transistors³⁰.

To alleviate the stringent atmospheric control that is needed to realize n-doped films with work functions of less than about 3.5 eV, the self-compensation strategy can be combined with precursor

n-dopants to generate the electron-doped state in the final device. As an illustration, a composite film of (4-(1,3-dimethyl-2,3-dihydro-1H-benzimidazol-2-yl)phenyl)dimethylamine (DMBI-H) dispersed in a naphthalene-bis(carboxyimide)-thiophene core with tethered trimethylammonium counter-ion (N4) yields strong electron-doping of the naphthalenebis(carboxyimide)-thiophene core when warmed (doping level of 1.0 electrons per repeat unit), whereas the corresponding composite film with the same naphthalene-bis(carboxyimide)-thiophene core but without the tethered counter-ion (N0) yields only weak doping (0.1 electrons per repeat unit) under identical conditions (Supplementary Methods and Extended Data Fig. 10). This reaffirms the role of the immobilized counter-ions in stabilizing doped states with extreme work functions.

Finally, we speculate about whether χ plays any part in the effective work function of these doped polymer heterojunctions where the polaron carriers co-exist with a high density of spectator ions. The V_{bi} measured in some diodes suggests that the effective work function can be favourably shifted by up to a few tenths of an electronvolt compared to the vacuum work function. For example, the n-doped N1 film in a DPPT2-TT diode (Fig 2c) exhibits an apparent decrease in effective work function by 0.3 eV, while a p-doped *para*-methoxy-substituted triarylamine-fluorene copolymer film in a PFOP diode exhibits an apparent increase in effective work function by 0.2 eV. These shifts point to a build-up of a favourable interfacial dipole at the heterojunction, perhaps by an ångström-scale displacement of the tethered counter-ions that, surprisingly, reinforces the extreme work function that is sought after in these films.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 13 January; accepted 3 October 2016.

- Groenendaal, L. B., Jonas, F., Freitag, D., Pielartzik, H. & Reynolds, J. R. Poly(3,4-ethylenedioxythiophene) and its derivatives: past, present, and future. *Adv. Mater.* **12**, 481–494 (2000).
- Belaine, D. *et al.* A high-performance p-doped conducting polymer blend based on sulfonated polyalkoxythiophene and poly(4-hydroxystyrene). *Chem. Mater.* **26**, 4724–4730 (2014).
- de Leeuw, D. M., Simenon, M. M. J., Brown, A. R. & Einerhand, R. E. F. Stability of n-type doped conducting polymers and consequences for polymeric microelectronic devices. *Synth. Met.* **87**, 53–59 (1997).
- Dai, A. *et al.* Investigation of p-dopant diffusion in polymer films and bulk heterojunctions: stable spatially-confined doping for all-solution processed solar cells. *Org. Electron.* **23**, 151–157 (2015).
- Hoven, C. V., Garcia, A., Bazan, G. C. & Nguyen, T. Q. Recent applications of conjugated polyelectrolytes in optoelectronic devices. *Adv. Mater.* **20**, 3793–3810 (2008).
- Huang, F., Wu, H. & Cao, Y. Water/alcohol soluble conjugated polymers as highly efficient electron transporting/injection layer in optoelectronic devices. *Chem. Soc. Rev.* **39**, 2500–2521 (2010).
- Lee, J. H. *et al.* Radical cation-anion coupling-induced work function tunability in anionic conjugated polyelectrolytes. *Adv. Energy Mater.* **5**, 1501292, (2015).
- Chan, H. S. O., Ho, P. K. H., Ng, S. C., Tan, B. T. G. & Tan, K. L. A new water-soluble, self-doping conducting polyaniline from poly(o-aminobenzylphosphonic acid) and its sodium salts: synthesis and characterization. *J. Am. Chem. Soc.* **117**, 8517–8523 (1995).
- Reilly, T. H., Hains, A. W., Chen, H. Y. & Gregg, B. A. A self-doping, O₂-stable, n-type interfacial layer for organic electronics. *Adv. Energy Mater.* **2**, 455–460 (2012).
- Lee, T. W. *et al.* Hole-injecting conducting-polymer compositions for highly efficient and stable organic light-emitting diodes. *Appl. Phys. Lett.* **87**, 231106 (2005).
- Zhou, Y. *et al.* A universal method to produce low-work function electrodes for organic electronics. *Science* **336**, 327–332 (2012).
- Wang, H. *et al.* Tuning contact recombination and open-circuit voltage in polymer solar cells via self-assembled monolayer adsorption at the organic-metal oxide interface. *J. Phys. Chem. C* **117**, 20474–20484 (2013).
- Belaine, D. *et al.* Perfluorinated ionomer-modified hole-injection layers: ultrahigh workfunction but non-ohmic contacts. *Adv. Funct. Mater.* **25**, 5504–5511 (2015).
- Tietze, M. L., Pahnner, P., Schmidt, K., Leo, K. & Lüssem, B. Doped organic semiconductors: trap-filling, impurity saturation, and reserve regimes. *Adv. Funct. Mater.* **25**, 2701–2707 (2015).
- Patil, A. O. *et al.* Self-doped conducting polymers. *Synth. Met.* **20**, 151–159 (1987).
- Walzer, K., Maennig, B., Pfeiffer, M. & Leo, K. Highly efficient organic devices based on electrically doped transport layers. *Chem. Rev.* **107**, 1233–1271 (2007).
- Lu, M., Nicolai, H. T., Wetzelaer, G. A. H. & Blom, P. W. M. n-Type doping of poly(p-phenylene vinylene) with air-stable dopants. *Appl. Phys. Lett.* **99**, 173302 (2011).
- Olthof, S. *et al.* Ultralow doping in organic semiconductors: evidence of trap filling. *Phys. Rev. Lett.* **109**, 176601 (2012).
- Zhou, M. *et al.* The role of δ -doped interfaces for Ohmic contacts to organic semiconductors. *Phys. Rev. Lett.* **103**, 036601 (2009).
- Kahn, A. Fermi level, work function and vacuum level. *Mater. Horiz.* **3**, 7–10 (2016).
- Png, R. Q. *et al.* Madelung and Hubbard interactions in polaron band model of doped organic semiconductors. *Nat. Commun.* **7**, 11948 (2016).
- Chia, P. J. *et al.* Direct evidence for the role of the Madelung potential in determining the work function of doped organic semiconductors. *Phys. Rev. Lett.* **102**, 096602 (2009).
- Chia, P. J. *et al.* Injection-induced de-doping in a conducting polymer during device operation: asymmetry in the hole injection and extraction rates. *Adv. Mater.* **19**, 4202–4207 (2007).
- Nicolai, H. T. *et al.* Unification of trap-limited electron transport in semiconducting polymers. *Nat. Mater.* **11**, 882–887 (2012).
- Bijleveld, J. C. *et al.* Poly(diketopyrrolopyrrole-terthiophene) for ambipolar logic and photovoltaics. *J. Am. Chem. Soc.* **131**, 16616–16617 (2009).
- Zou, Y. *et al.* A thieno[3,4-c]pyrrole-4,6-dione-based copolymer for efficient solar cells. *J. Am. Chem. Soc.* **132**, 5330–5331 (2010).
- Cheng, X. *et al.* Controlling electron and hole charge injection in ambipolar organic field-effect transistors by self-assembled monolayers. *Adv. Funct. Mater.* **19**, 2407–2415 (2009).
- Brondijk, J. J., Torricelli, F., Smits, E. C. P., Blom, P. W. M. & de Leeuw, D. M. Gate-bias assisted charge injection in organic field-effect transistors. *Org. Electron.* **13**, 1526–1531 (2012).
- Allain, A., Kang, J., Banerjee, K. & Kis, A. Electrical contacts to two-dimensional semiconductors. *Nat. Mater.* **14**, 1195–1205 (2015).
- Zaumseil, J., Friend, R. H. & Sirringhaus, H. Spatial control of the recombination zone in an ambipolar light-emitting organic transistor. *Nat. Mater.* **5**, 69–74 (2006).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank C. Hu, H. Guo, D. Belaine, X.-Y. Hou, M.-H. Teo, K. Yeo, S.-N. Tan, Z.-W. Tan, J. Chen, K. Lim, Y.-S. Soh, P. Tanay and S.-C. Lee for contributions to the experimental work. We thank I. Grizzi, R. Wilson and the CDT/Sumitomo Chemical Co. team for materials and support. We particularly thank R. H. Friend for inspiration and insights. This research is partially supported by the National Research Foundation, Prime Minister's Office, Singapore under its Competitive Research Programme (CRP Award No. NRF-CRP 11-2012-03: R-144-000-339-281, R-143-000-608-281), and the Ministry of Education, Singapore (R-144-000-324-112). Solar Energy Research Institute of Singapore (SERIS) is sponsored by the National University of Singapore and Singapore's National Research Foundation through the Singapore Economic Development Board.

Author Contributions R.-Q.P. led the heterostructure and interface work, L.-L.C. led the materials chemistry work, and P.K.H.H. led the device physics work. C.G.T., K.-K.C. and M.C.Y.A. fabricated and characterized devices. K.-K.C., M.C.Y.A., V.K. and M.N.S. synthesized materials. J.-K.T. simulated device behaviour. J.H.B. and T.K. contributed directions. All authors discussed the experiments and results. R.-Q.P., L.-L.C. and P.K.H.H. wrote the manuscript.

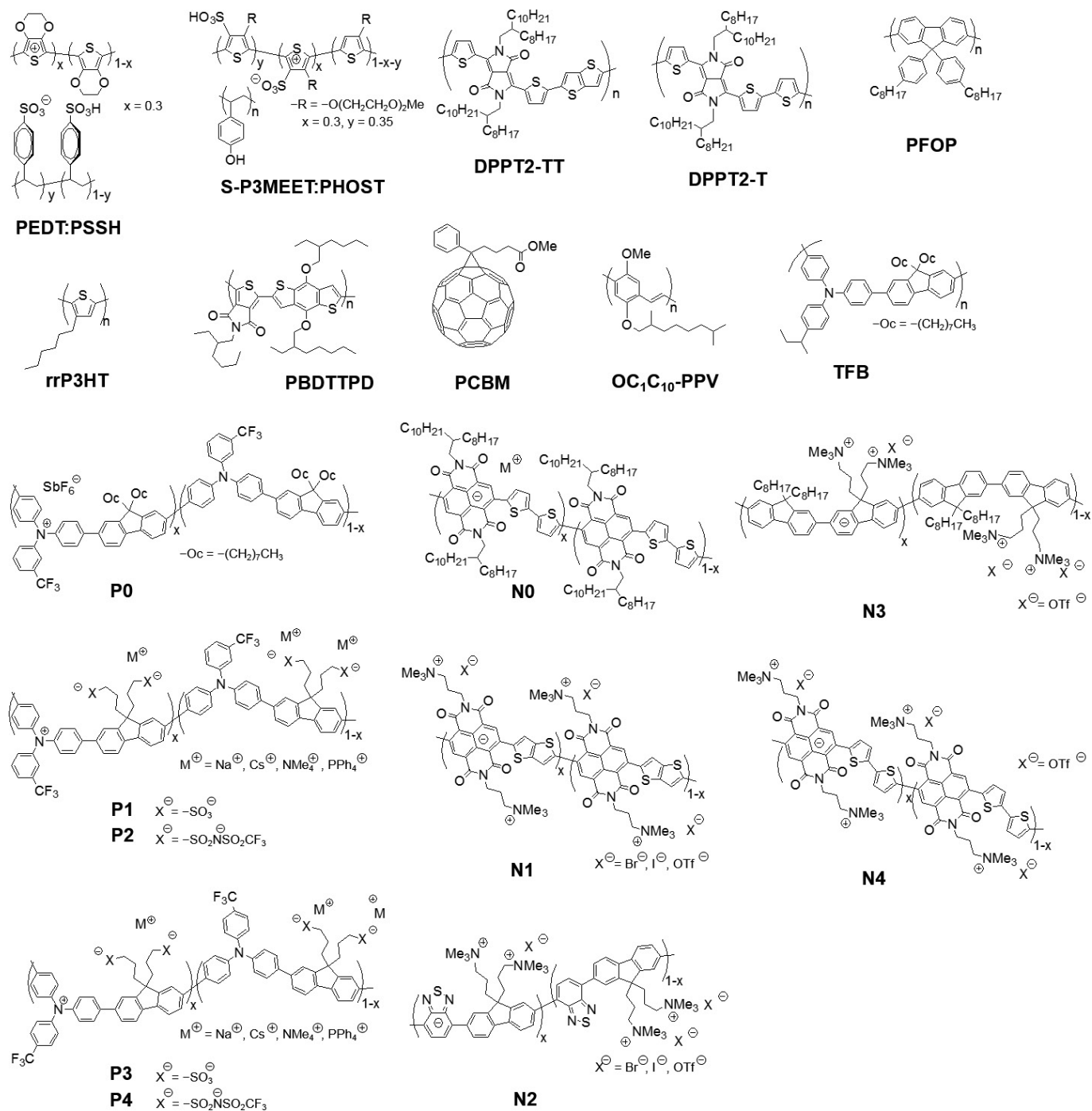
Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.K.H.H. (phyhop@nus.edu.sg), R.-Q.P. (phypngqr@nus.edu.sg) or L.-L.C. (chmcll@nus.edu.sg).

Reviewer Information Nature thanks A. Facchetti, D. Seferos and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

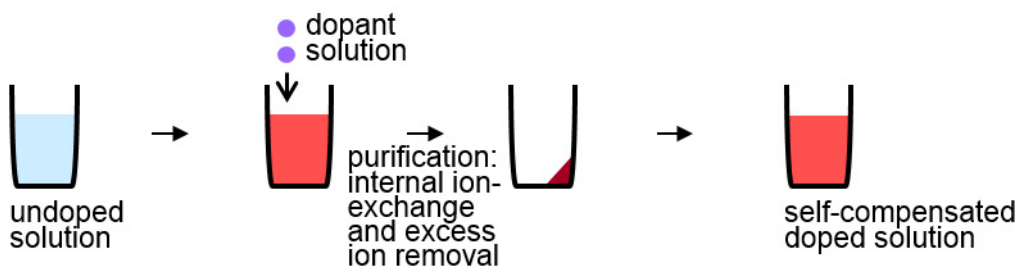
Data availability. The data supporting the findings of this study are available within this Letter, its Extended Data and its Supplementary Information.

31. Sawyer, D. T., Sobkowiak, A. & Roberts, J. L. *Electrochemistry for chemists* 2nd edn, 358–402 (Wiley, 1995).
32. Wood, P. M. The potential diagram for oxygen at pH 7. *Biochem. J.* **253**, 287–289 (1988).
33. Zhou, M. *et al.* Effective work functions for the evaporated metal/organic semiconductor contacts from *in-situ* diode flatband potential measurements. *Appl. Phys. Lett.* **101**, 013501 (2012).
34. Chua, L. L. *et al.* General observation of n-type field-effect behaviour in organic semiconductors. *Nature* **434**, 194–199 (2005).
35. Wei, P., Oh, J. H., Dong, G. & Bao, Z. Use of a 1*H*-benzoimidazole derivative as an *n*-type dopant and to enable air-stable solution-processed *n*-channel organic thin-film transistors. *J. Am. Chem. Soc.* **132**, 8852–8853 (2010).

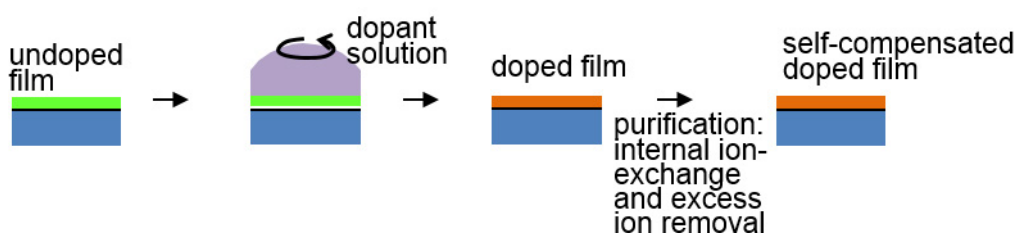


Extended Data Figure 1 | Chemical structures of materials used. The value of x denotes the doping level, in carriers per repeat unit (r.u.), with $x = 0$ denoting an undoped polymer and $x = 1$ a fully doped polymer.

Solution-state self-compensated doping:

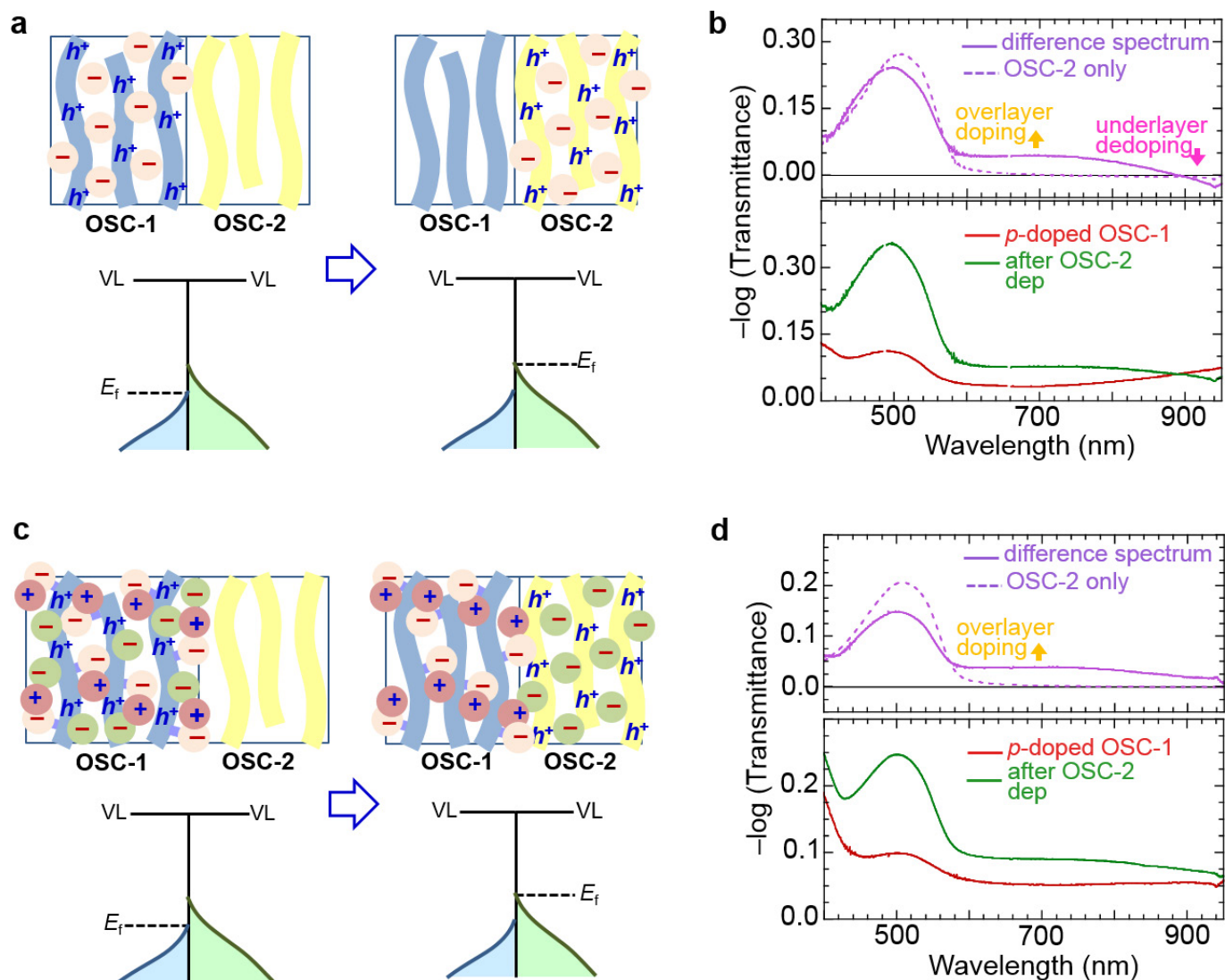


Film-state self-compensated doping:



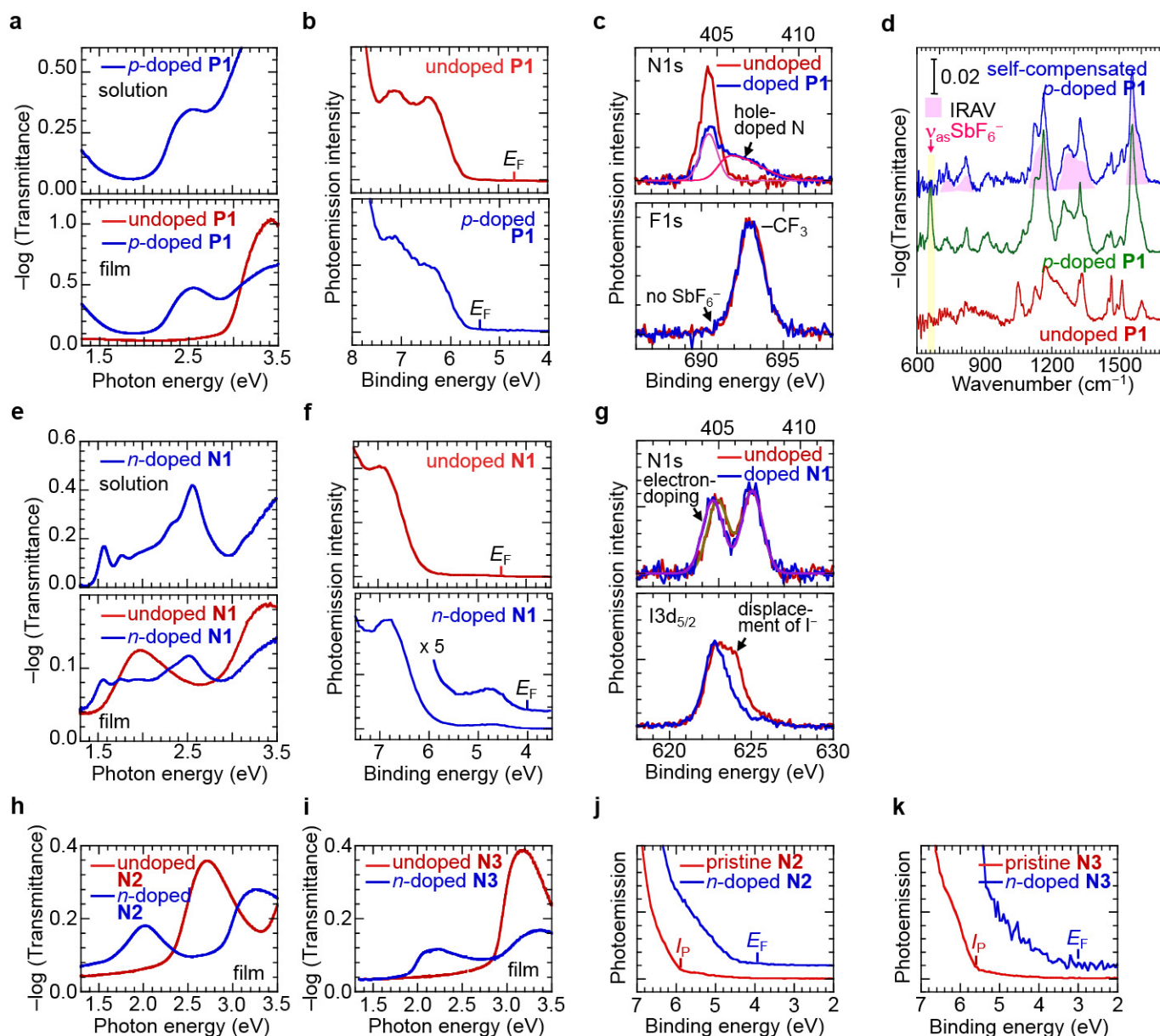
Extended Data Figure 2 | Schematic of the preparation of self-compensated heavily doped polymer organic semiconductors. Top, solution-state doping method. The undoped polymer dissolved in a suitable solvent is doped in solution by the desired dopant, and purified by repeated precipitation with a non-solvent while keeping the unwanted ions in solution. This process yields the fully self-compensated doped

polymer as a solid, which can be dissolved in the desired solvent. Bottom, film-state doping method. The undoped polymer film is deposited on the desired substrate, doped by contact with a dopant solution, and purified by repeated contact with a non-solvent that dissolves the unwanted ions. This directly yields the fully self-compensated doped polymer film.



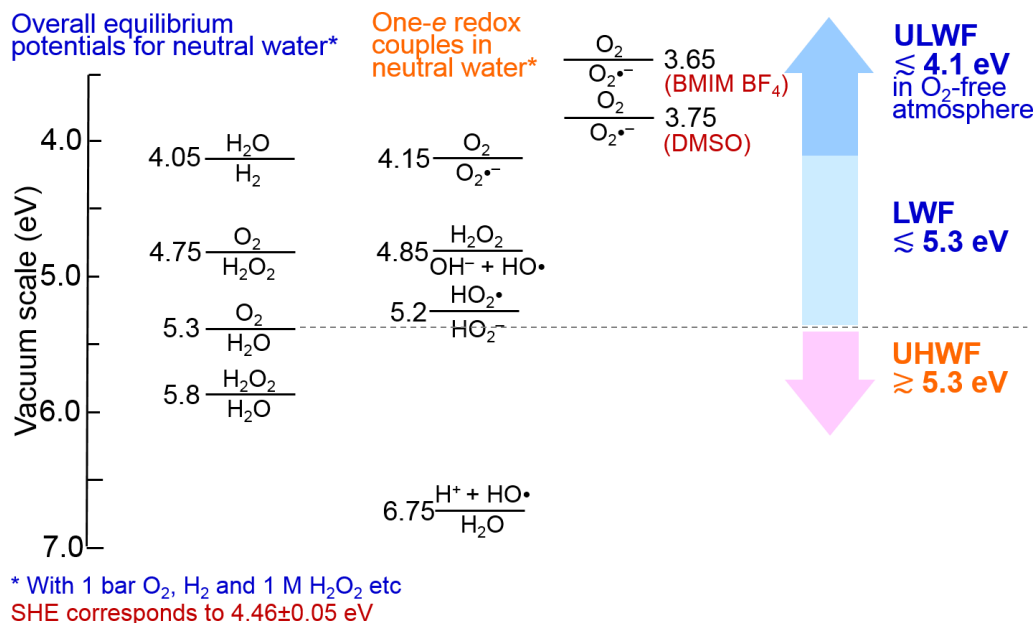
Extended Data Figure 3 | Doping-profile migration in the presence of mobile counter-ions. **a**, **b**, Schematic (a) and experimental data (b) for transfer of holes and counter-ions from a hole-doped organic semiconductor layer with mobile counter-anions (OSC-1) to an adjacent organic semiconductor layer (OSC-2) during solution processing. OSC-1 is **P0** ($I_p = 5.9$ eV), hole-doped to $x \approx 0.8$, with SbF_6^- as counter-ion. OSC-2 is OC_1C_{10} -PPV ($I_p = 5.0$ eV), deposited by spin-casting on p-doped **P0**. **c**, **d**, Schematic (c) and experimental data (d) for transfer of holes and counter-ions from an incompletely self-compensated hole-doped organic semiconductor (OSC-1) to an adjacent organic

semiconductor layer (OSC-2) during solution processing. OSC-1 is incompletely self-compensated hole-doped **P1** ($x \approx 0.8$) containing 50% residual Na^+ and SbF_6^- . OSC-2 as in **a** and **b**. Nearly complete transfer of holes and counter-anions from the doped underlayer to the OC_1C_{10} -PPV overlayer occurred during spin-casting, as evidenced by emergence of the polaron band of OC_1C_{10} -PPV (600–800 nm) and bleaching of the polaron band of **P0** and **P1** (>900 nm). This doping transfer is driven by thermodynamics from a high- I_p layer to a low- I_p layer. E_F , Fermi energy; VL, vacuum level.



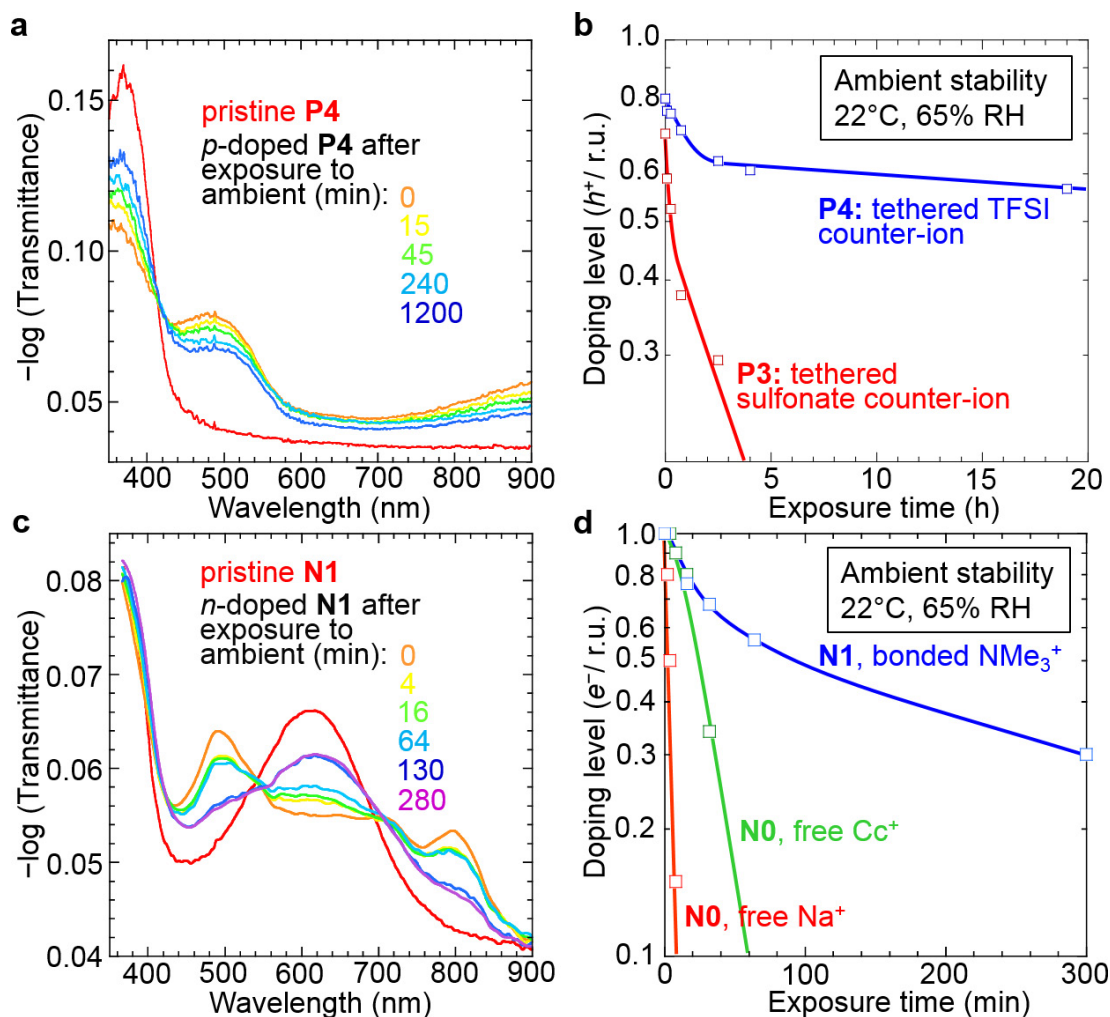
Extended Data Figure 4 | Characterization of selected self-compensated, heavily doped polymers. **a–d**, Triarylamine-fluorene copolymer **P1**. **a**, Optical spectra of doped **P1** in solution (top), and of a **P1** film before and after doping (bottom). For solution-state doping, 0.12 mM r.u. **P1** in propylene carbonate was doped with NOSbF_6 and self-compensated, yielding a final doping level of 0.5 holes per r.u. The optical spectrum is collected in a cell with a 2.0-mm path length. For film-state doping, a 100-nm-thick **P1** film was doped with NOSbF_6 and self-compensated, yielding a final doping level of 0.9 holes per r.u. **b**, **c**, UPS (**b**) and XPS spectra (**c**; N1s and F1s) of undoped and doped **P1** ($x \approx 0.5$) films. **d**, FTIR evidence for full self-compensation in a **P1** film. Spectra were acquired before doping (red), after film-state doping with 6 mM NOSbF_6 in acetonitrile (green) and after self-compensation by excess ion removal (Na^+ and SbF_6^-) via spin-rinse with acetonitrile (blue). Infrared active vibration modes (IRAV; 1,100–1,600 cm^{-1}) indicate that the film is doped. SbF_6^- (anti-symmetric stretching frequency $\nu_{\text{as}} = 660 \text{ cm}^{-1}$) is present before, but not after, self-compensation (residual, <1%).

e–g, Naphthalenetetracarboxydiimide-thiophene copolymer **N1**. **e**, Optical spectra of doped **N1** in solution (top), and of an **N1** film before and after doping (bottom). For, solution-state doping, 0.22 mM r.u. **N1** in dimethylsulfoxide was doped with sodium 9,10-diphenylanthracene and self-compensated, yielding a final doping level of 1.0 electrons per r.u. The optical spectrum is collected in a cell with a 2.0-mm path length. For film-state doping, a 25-nm-thick **N1** film was doped with sodium 9,10-diphenylanthracene and self-compensated, yielding a final doping level of 1.0 electrons per r.u. **f**, **g**, UPS (**f**) and XPS spectra (**g**; N1s and $\text{I3d}_{5/2}$) of undoped and doped **N1** ($x \approx 1.0$) films. **h–k**, Fluorene copolymers **N2** and **N3**. **h**, **i**, Optical spectra of 40-nm-thick **N2** (**h**) and **N3** (**i**) films, before and after film-state doping with sodium naphthalenide solution and spin-rinsing with tetrahydrofuran. Final doping levels, 0.9 electrons per r.u. (**N2**; **h**) and 0.4 electrons per r.u. (**N3**; **i**). **j**, **k**, UPS spectra of undoped and n-doped **N2** (**j**) and **N3** (**k**) films. All binding energies are referenced to vacuum level. The Fermi energy E_F is marked for doped films; the ionization potential I_P is marked for undoped films.



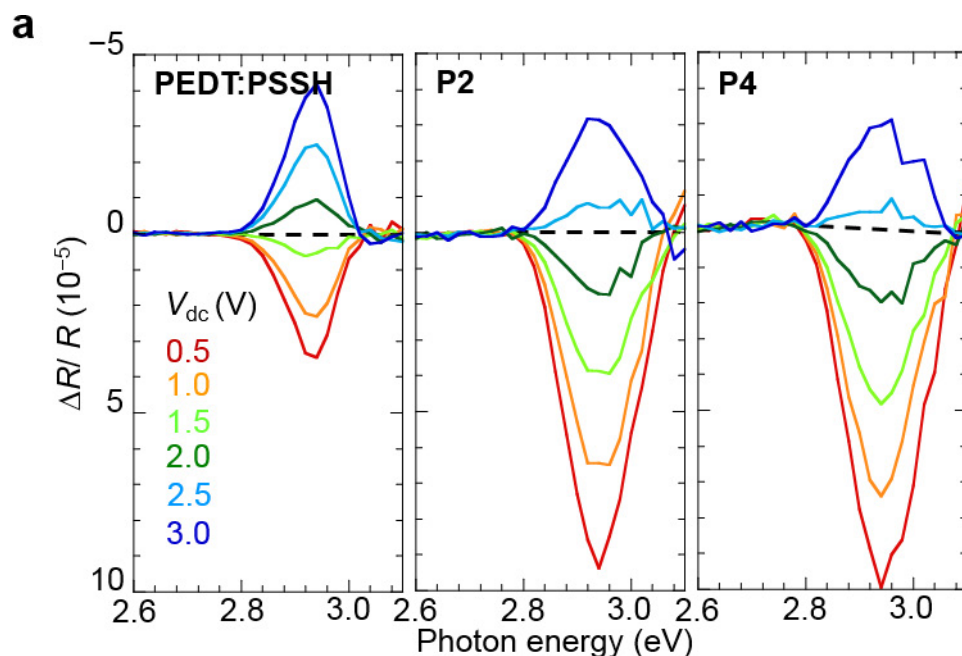
Extended Data Figure 5 | Electrode potentials of H₂O and O₂ redox couples in pH 7 water, plotted on the vacuum scale. Conditions: neutral H₂O; pressure, 1 bar (O₂, H₂); concentration, 1 M (H₂O₂, O₂^{•−}, HO[•], HO₂[•]). The reference voltage (0.00 V) on the standard hydrogen electrode (SHE) scale corresponds to 4.46 eV on the vacuum scale. Electrode potentials of O₂/O₂^{•−} were computed from ref. 31; other electrode potentials were computed from ref. 32. p-doped films with ultrahigh work functions (UHWF) of greater than about 5.3 eV and n-doped films with low work functions (LWF) of less than about 5.3 eV can oxidize neutral H₂O to O₂, and reduce O₂ to neutral H₂O, respectively, and so cannot remain indefinitely stable in the presence of these species. If the oxygen reduction reaction proceeds through one-electron steps, then

the relevant first one-electron couple (O₂/O₂^{•−}) lies in the range 4.15–3.65 eV, depending on the dielectric medium, which opens up a stability window for films with ultralow work functions (ULWF). The stability of the films with ultralow work functions is ultimately limited by the electrode potential at the minimum practical O₂ concentration during film preparation (p.p.m. range). To overcome this limitation, precursor n-dopants can be incorporated into the self-compensated platform (see main text). Conversely, the relevant first one-electron couple for water oxidation ((H⁺ + HO[•])/H₂O) lies at 6.75 eV, which opens up a stability window for films with ultrahigh work functions. DMSO, dimethyl sulfoxide; BMIM BF₄, 1-butyl-3-methylimidazolium tetrafluoroborate.

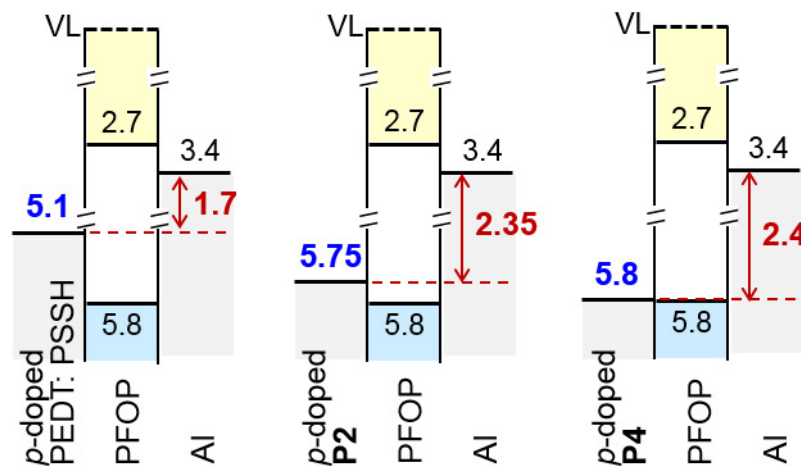


Extended Data Figure 6 | Ambient stability of ultrahigh- and ultralow-work-function films. **a, b,** Self-compensated p-doped P4. **a,** Time-dependent optical spectra of a 12-nm-thick film in ambient conditions (295 K, 65% relative humidity (RH)). The initial doping level was $x \approx 0.8$. **b,** To compare the stability imparted by different counter-anions, the doping level is shown as a function of exposure time for P4 (with tethered trifluoromethylsulfonylimide counter-anion; blue) and P3 (with tethered sulfonate counter-anion; red). The films were hole-doped with 1 mM NOSbF_6 in acetonitrile and self-compensated by spin-rinse with acetonitrile in a N_2 glove box, and then exposed to ambient conditions and measured after specified intervals. The doping level was evaluated from the integrated polaron band intensity (460–540 nm). Time to 50% doping level: P4, 50 h; P3, 15 min from initial doping level of about 0.8 holes per r.u. **c, d,** Self-compensated n-doped N1. **c,** Time-dependent optical spectra

of a 7.5-nm-thick self-compensated film in ambient conditions. The initial doping level was $x \approx 1.0$. **d,** To compare the stability of self-compensated and non-self-compensated films, the doping level is shown as a function of exposure time for N1 (with tethered trimethylammonium counter-ion; blue) and N0 with unbonded Na^+ (red) or Cc^+ (cobaltocenium; green) counter-ions. The films were electron doped with 15 mM sodium 9,10-diphenylanthracene in diglyme and self-compensated by spin-rinse with acetonitrile. The doping level was evaluated from the integrated band intensity (540–700 nm) after calibration. Time to 50% doping level: N1, 2 h; N0 counter-balanced by Cc^+ , 20 min; N0 counter-balanced by Na^+ , 4 min from initial doping level of about 1.0 electrons per r.u. Effective ionic radius r_{ion} : Na^+ , 1.0 Å; Cc^+ , 3. Å; $-(\text{CH}_2)_3\text{NMe}_3^+$, 3.2 Å. De-doping recovers the original π – π^* spectrum.

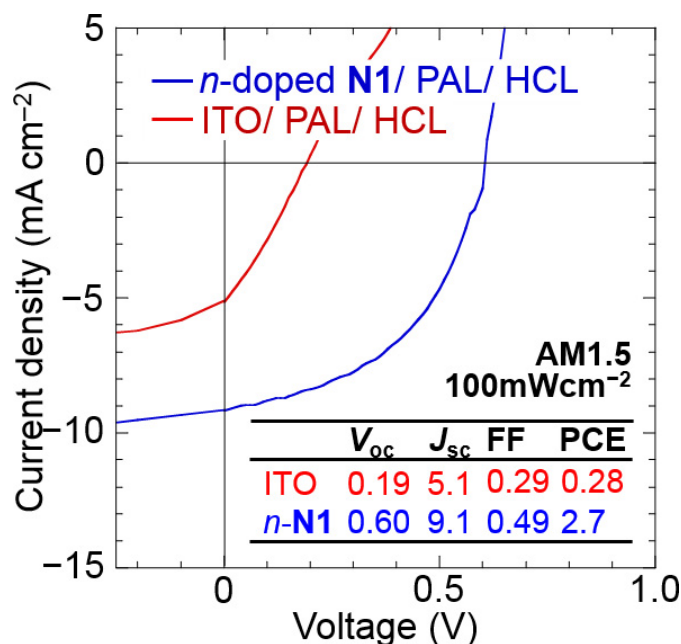


b Apparent energy-level alignment diagrams

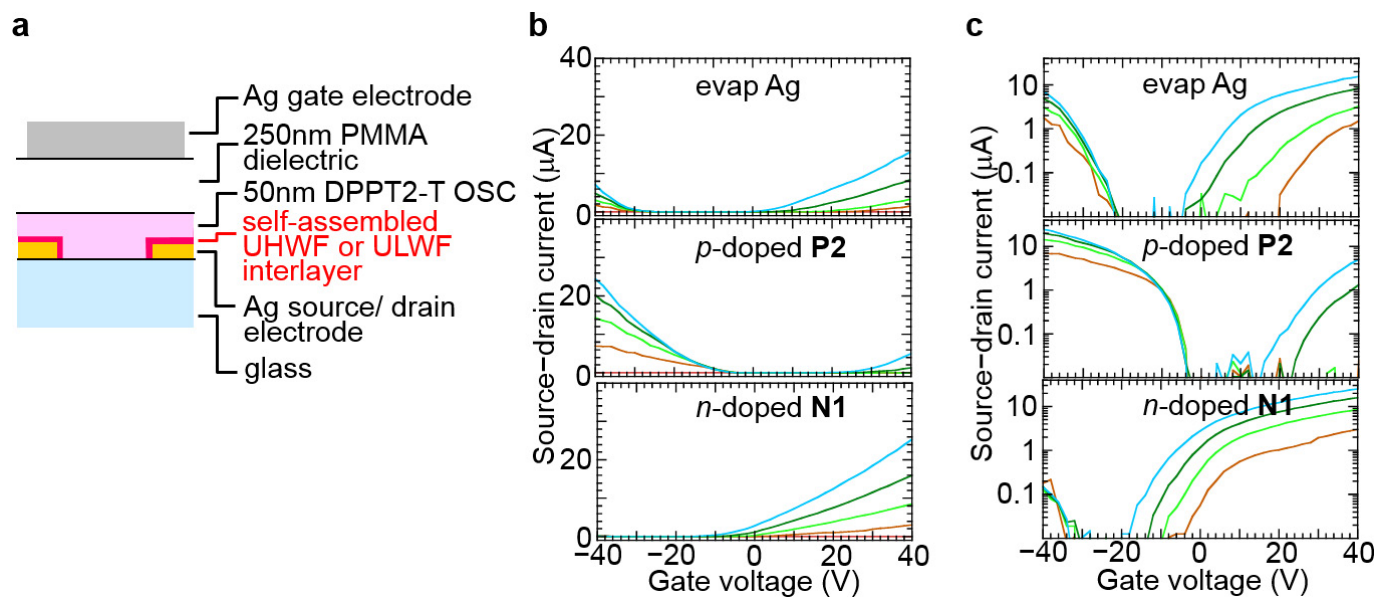


Extended Data Figure 7 | Apparent work function evaluated by *in situ* built-in potential measurements. **a**, Stark-effect electro-absorption spectra of hole-only diodes with p-doped PEDT:PSSH (left), **P2** (middle) and **P4** (right) as hole contacts, PFOP as the organic semiconductor and Al as the electron-blocking contact. Normalized modulated reflection spectra ($\Delta R/R$) are shown for different d.c. biases V_{dc} applied to the hole contact, where ΔR is the modulated root-mean-square reflectance in-phase with the forward-bias half-cycle and R is the d.c. reflectance³³. Negative $\Delta R/R$ therefore indicates induced absorption in-phase with the forward-bias half-cycle. The dotted line gives the null background. Temperature, 30 K; modulation frequency, 535 Hz; root-mean-square voltage V_{rms} , 0.5 V. **b**, Apparent energy-level alignment diagrams deduced from the built-in

potential V_{bi} , assuming rigid alignment at the electron contact. Energies are plotted relative to the vacuum level (VL) of the organic semiconductor. The Stark band at 2.87 eV near the π - π^* absorption edge of PFOP shows polarity inversion at 1.70 V for PEDT:PSSH, at 2.35 V for **P2** and at 2.4 V for **P4**. Estimated uncertainty, ± 0.03 V. Consequently, the apparent work functions are: PEDT:PSSH, $3.4 + 1.7 = 5.1$ eV; **P2**, $3.4 + 2.35 = 5.75$ eV; **P4**, $3.4 + 2.4 = 5.8$ eV. In reality, the effective work function at the electron contact decreases in response to an increase in work function at the hole contact, driving it into the pinned regime. This then gives rise to pinning at about 5.6 eV. The apparent work functions therefore confirm that the work functions of **P2** and **P4** are both larger than 5.6 eV, with the work function of **P4** greater than that of **P2**.

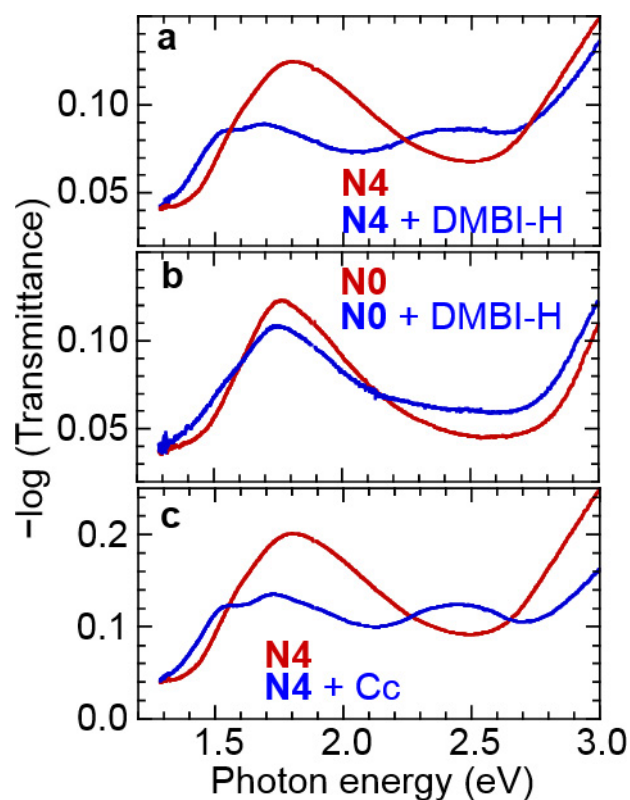


Extended Data Figure 8 | Characteristics of solar cells with an ultralow-work-function polymer as the bottom electron-collection contact. Current density versus voltage for an 80-nm-thick P3HT:PCBM photoactive layer (weight ratio of 1: 0.8) under 100 mW cm^{-2} illumination (spectral-mismatch corrected) with an ultralow-work-function polymer as the bottom electron-collection contact and PEDT:PSSH as the top hole-collection contact. 20-nm-thick n-doped N1 was deposited as the ECL via spin-coating from DMSO solution onto the ITO layer. The photoactive layer (PAL) was deposited from chlorobenzene. 30-nm-thick PEDT:PSSH was deposited as the HCL via spin-coating from aqueous solution onto the top of the PAL and capped with evaporated Ag. A control cell was fabricated in the same batch by omitting the N1 layer, so that ITO provided the electron-collection contact. The control cell (red) shows poor behaviour with an S-shaped J - V characteristic, whereas the test cell (blue) gives good performance similar to standard cells with evaporated Ca as the electron-collection contact. Open-circuit voltage V_{oc} is given in V; short-circuit current J_{sc} in mA cm^{-2} , and power conversion efficiency PCE in %. FF, fill factor. Both cells have been corrected for shunt leakage.



Extended Data Figure 9 | Facile electrode differentiation by self-aligned assembly of ultrahigh- and ultralow-work-function interlayers on Ag. **a**, Schematic of the FET structure. **b**, **c**, Transfer characteristics on a linear (**b**) and semi-logarithmic (**c**) scales. A monolayer of p-doped P2 or n-doped N1 was self-assembled onto the Ag electrode array from dilute

solution. Channel length, 100 μm ; channel width, 3 mm; V_d was stepped from 0 V (red) or to -20 V (blue) in -5 V steps. The control FET shows the expected ambipolar behaviour³⁴, whereas the modified electrodes show gate-source V_{gs} threshold shifts, similar to that observed on Au electrodes.



Extended Data Figure 10 | Integration of precursor n-dopant methodology into self-compensated doped polymers. **a, b,** N4 (**a**) and N0 (**b**) with (blue) and without (red) DMBI-H. Pristine polymer films (15-nm-thick N4 and 20-nm-thick N0) and composite films with DMBI-H (1:2 mol/mol r.u.) were spin-cast and baked at 120 °C on a hotplate, and then measured by transmission optical spectroscopy. DMBI-H is (4-(1,3-dimethyl-2,3-dihydro-1H-benzimidazol-2-yl)phenyl) dimethylamine, a precursor n-dopant³⁵. **c,** Authentic n-doping of N4 films, as a reference. A 30-nm-thick N4 film was spin-cast and then electron-doped to 1.0 electron per r.u. by contact with cobaltocene (Cc; 2 mM in diglyme). N0 has an identical π -conjugated organic semiconductor core to N4, but lacks tethered counter-ions. Whereas N4 exhibits strong doping, N0 exhibits only weak doping under identical conditions, possibly owing to the lower stability of its n-doped state.

The evolving quality of frictional contact with graphene

Suzhi Li^{1,2,3}, Qunyang Li^{4,5}, Robert W. Carpick⁶, Peter Gumbsch^{7,8}, Xin Z. Liu⁶, Xiangdong Ding¹, Jun Sun^{1,9} & Ju Li^{1,2,9}

Graphite and other lamellar materials are used as dry lubricants for macroscale metallic sliding components and high-pressure contacts. It has been shown experimentally that monolayer graphene exhibits higher friction than multilayer graphene and graphite, and that this friction increases with continued sliding, but the mechanism behind this remains subject to debate. It has long been conjectured that the true contact area between two rough bodies controls interfacial friction¹. The true contact area, defined for example by the number of atoms within the range of interatomic forces, is difficult to visualize directly but characterizes the quantity of contact. However, there is emerging evidence that, for a given pair of materials, the quality of the contact can change, and that this can also strongly affect interfacial friction^{2–7}. Recently, it has been found that the frictional behaviour of two-dimensional materials exhibits traits^{8–13} unlike those of conventional bulk materials. This includes the abovementioned finding that for few-layer two-dimensional materials the static friction force gradually strengthens for a few initial atomic periods before reaching a constant value. Such transient behaviour, and the associated enhancement of steady-state friction, diminishes as the number of two-dimensional layers increases, and was observed only when the two-dimensional material was loosely adhering to a substrate⁸. This layer-dependent transient phenomenon has not been captured by any simulations^{14,15}. Here, using atomistic simulations, we reproduce the experimental observations of layer-dependent friction and transient frictional strengthening on graphene. Atomic force analysis reveals that the evolution of static friction is a manifestation of the natural tendency for thinner and less-constrained graphene to re-adjust its configuration as a direct consequence of its greater flexibility. That is, the tip atoms become more strongly pinned, and show greater synchrony in their stick-slip behaviour. While the quantity of atomic-scale contacts (true contact area) evolves, the quality (in this case, the local pinning state of individual atoms and the overall commensurability) also evolves in frictional sliding on graphene. Moreover, the effects can be tuned by pre-wrinkling. The evolving contact quality is critical for explaining the time-dependent friction of configurationally flexible interfaces.

We conducted molecular dynamics simulations by sliding a silicon tip over graphene supported on an amorphous Si (a-Si) substrate (see Methods). Substrate-graphene adhesion was modelled as a van der Waals interaction with an effective work of adhesion of approximately 0.1 J m^{-2} , based on experimental values¹⁶. Figure 1a shows the relaxed system consisting of a-Si and a monolayer of graphene at 300 K. As in experiments¹⁶, the modelled surface morphology of graphene conformed well to the a-Si substrate (Fig. 1b and c). A rigid round

(001)-oriented crystalline silicon tip with a radius of 16.3 nm was placed in contact with graphene (Fig. 1d) and equilibrated for 2 ns (Fig. 1e and f). Graphene's high out-of-plane flexibility coupled with tip-graphene adhesion causes local puckering near the contact edge (Fig. 1e, arrow). A similar puckering phenomenon was also found for the multilayer graphene systems (see Supplementary Discussion 1).

We then applied an external normal load of 0.8 nN to the tip and further relaxed the whole system. After relaxation, we displaced a harmonic spring laterally coupled to the tip at 2 m s^{-1} . It is noted that the normal load was stably maintained at 0.8 nN during the friction simulation (see Supplementary Discussion 2). As the tip slid on graphene (with layer numbers $N = 1–4$), clear stick-slip motion with two distinct stages was observed in all cases (Fig. 1g), in contrast to conventional atomic stick-slip on bulk graphite¹⁷. Initially, the local peaks of the lateral force (that is, the static friction force, when slip initiates) increased progressively with each slip. After a few periods, the behaviour became regular, and the peak forces remained constant. The slip distance in each stick-slip event is approximately 2.5 \AA , equivalent to the lattice spacing of graphene. The simulations consistently reproduce, for the first time, the two-stage friction behaviour observed in experiments⁸.

Our simulations also show layer-dependent friction⁸, that is, the strengthening effect became weaker and the frictional force reduced as sample thickness increased; the strengthening almost vanished once the thickness reached four layers. To check that two-stage friction was not a dynamic effect caused by the high sliding speed, we first reduced the speed to 1 m s^{-1} and 0.5 m s^{-1} and found that there is only a weak dependence of the frictional strengthening on sliding speed (see Supplementary Discussion 3). In addition, we performed quasi-static calculations^{5,18} and found that the two-stage friction trait was retained (see Supplementary Discussion 4). Finally, we computed the energy barriers for stick-slip motion, and confirmed that the stick-slip event is not driven primarily by thermal fluctuations at experimental time and force scales, but primarily by the external forcing (see Supplementary Discussion 5).

Early studies attributed the friction enhancement for thinner two-dimensional samples to puckering^{8,14,19}, where the sliding tip induces more out-of-plane deformation for thinner samples, leading to a larger true contact area and thereby a larger friction. Here, we also observed noticeable puckering (Fig. 1e, Supplementary Video) and found that the puckered configuration evolved as the tip moved forward. Figure 1h shows the average friction force together with the contact area in the constant static friction force regime (that is, beyond the initial strengthening regime) for different layered graphene samples. The simulations show that a larger contact area corresponded to larger friction, and was correlated with increased puckering. However, the friction increase

¹State Key Laboratory for Mechanical Behavior of Materials and Frontier Institute of Science and Technology, Xi'an Jiaotong University, Xi'an 710049, China. ²Department of Nuclear Science and Engineering and Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. ³Institute of Nanotechnology, Karlsruhe Institute of Technology, 76344 Eggenstein-Leopoldshafen, Germany. ⁴Center for Nano and Micro Mechanics, Applied Mechanics Laboratory, School of Aerospace Engineering, Tsinghua University, Beijing 100084, China. ⁵State Key Laboratory of Tribology, Tsinghua University, Beijing 100084, China. ⁶Department of Mechanical Engineering and Applied Mechanics, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. ⁷Institute for Applied Materials, Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany. ⁸Fraunhofer Institute for Mechanics of Materials IWM, 79108 Freiburg, Germany. ⁹Center for Advancing Materials Performance from the Nanoscale (CAMP-Nano), School of Materials Science and Engineering, Xi'an Jiaotong University, Xi'an 710049, China.

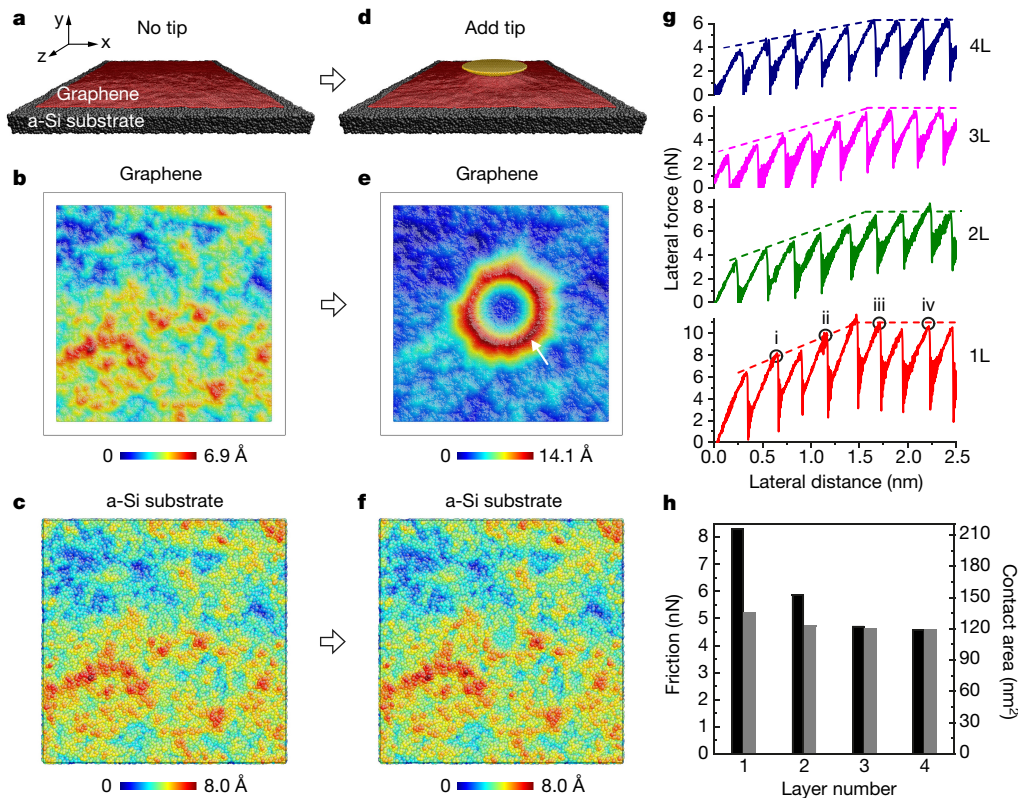


Figure 1 | Model setup and frictional behaviour for a Si tip sliding over a graphene/a-Si substrate system at 300 K. **a**, Graphene adhering to the a-Si substrate. **b**, **c**, Surface morphology of monolayer graphene (**b**) and the substrate (**c**). **d**, A rigid Si tip is placed in contact with graphene for the friction tests. **e**, **f**, Morphology of monolayer graphene (**e**) and substrate (**f**) in the presence of the tip after 2 ns. The lateral dimensions of the substrate and graphene are 43.4 nm × 43.4 nm and 38.5 nm × 38.6 nm, respectively. **g**, Force traces showing stick-slip behaviour on single (1L) and multilayer

(an 80% increase for monolayer versus four-layer graphene, consistent with experiments) was substantially larger than the change in contact area (14%). Therefore, the true contact area (which by definition is geometric and additive) change caused by the puckering of graphene alone cannot account for the large increase in friction. This strongly suggests the existence of additional mechanisms contributing to the layer dependence of friction and the transient strengthening.

To trace the physical processes at play, we analysed the distribution of atomistic interaction forces at the interface. For the tip, the net lateral interaction force from graphene, F_{interact} , is the sum of the x components of all atomic forces exerted by graphene, $F_{\text{interact}} = \sum_i f_i^{\text{friction}} = \sum_i (f_i(x_{\text{spr}}) - f_i(0))$, where $f_i(x_{\text{spr}})$ is the x component of the graphene–tip interaction force for tip atom i when the spring moves a distance x_{spr} , and $f_i(0)$ is the interaction force before sliding. To highlight the resultant forces due to sliding, we considered only the incremental interfacial forces with respect to the initial state (see Supplementary Discussion 6).

We first performed interfacial force analysis for monolayer graphene. Figure 2a–d shows the f_i^{friction} distributions at four moments, each when the lateral force reached its local peak values during the stick-slip friction, marked i, ii, iii and iv in Fig. 1g. Owing to the crystallographic and geometrical differences between graphene and silicon, f_i^{friction} values at all four moments exhibited a seemingly random distribution; the total sum in each case was negative (thus producing frictional resistance). The f_i^{friction} distribution clearly evolved from points i to iii (Fig. 2a–c) and was almost unchanged between points iii and iv (Fig. 2c and d).

We individually followed f_i^{friction} of a few tip atoms that were acting as pinning sites or ‘traps’ (see Supplementary Discussion 7). For some of these atoms, the absolute value of f_i^{friction} increased in the strengthening

(2L–4L) graphene/a-Si substrates. **h**, Variations of averaged friction (black) and contact area (grey) with number of layers of graphene. The black, red and gold atoms in **a** and **d** refer to the a-Si substrate, graphene and the tip, respectively; atoms in **b**, **c**, **e** and **f** are coloured according to the height amplitude along the y direction. The tip–graphene contact area is taken to be ms , where m is the number of graphene atoms that are in intimate contact with the tip atoms and s (2.77 \AA^2 per atom) is the atomic area of graphene (see Methods for details).

stage and then saturated. This indicates that pinning gets progressively stronger at these sites with each slip. These atoms apparently provide a substantial contribution to frictional strengthening. For most atoms, the variation of f_i^{friction} was steady, indicating no pinning enhancement. These atoms contribute to the overall friction, but not to the initial transient strengthening by local trap deepening.

Further inspection shows that the variation of the interfacial force is closely related to the local configuration of the contacting graphene. If the environment near a tip atom allows the graphene to locally adjust its atomic configuration, the interfacial force will exhibit an evolution effect. However, the detailed local contact environment varies from atom to atom owing to complex tip, graphene and rough substrate configurations. We cannot tell beforehand how an individual tip atom and its surrounding would evolve, because they not only depend on the initial configuration but also depend on the randomly rough surface and the complicated deformation history (for example, the statistical features of the final configuration depend on the entire sliding path, and not just on the initial and final tip locations). Nevertheless, our simulations clearly demonstrate that the interface tends to evolve towards a more commensurate (that is, with more atomic alignment at the interface) and better-pinned state (that is, it finds a deeper energy trap).

To quantify this, we plotted the histogram of the atomic interaction forces at these four moments in Fig. 2e. All four histograms have larger populations at low-magnitude interaction forces and much lower populations at high magnitude. However, these histograms were all skewed negatively because the overall net force was negative. More importantly, as indicated by the insets of Fig. 2e, the range of the distribution increased appreciably from points i to iv (for example, the maximum forces are around $\pm 0.6 \text{ nN}$ at point i, but increase to $\pm 0.8 \text{ nN}$

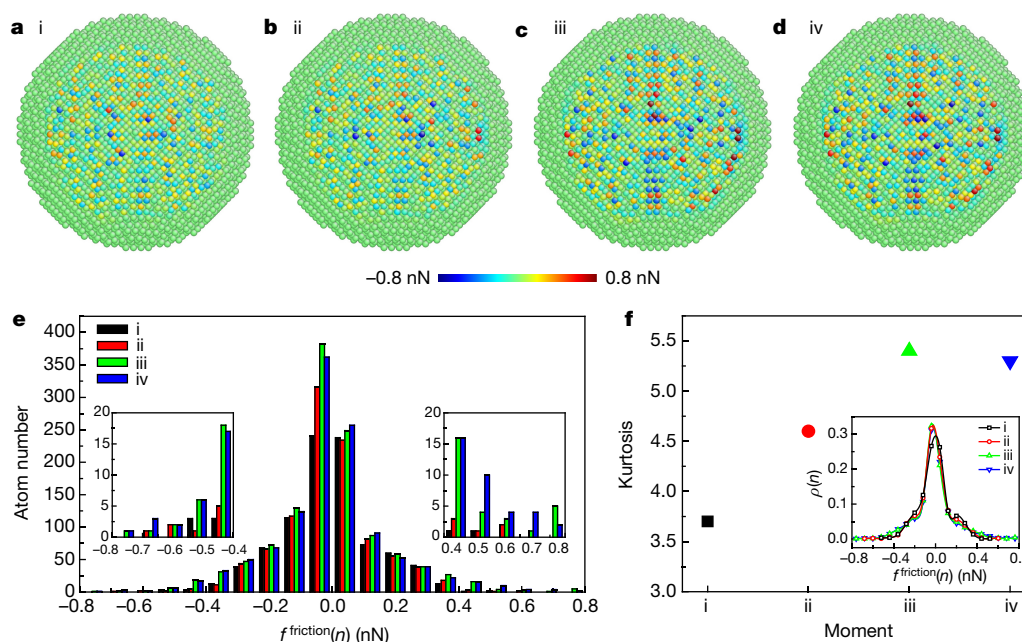


Figure 2 | Evolution of the atomic-level forces contributing to friction on a monolayer graphene/a-Si substrate. a–d, Corresponding f_i^{friction} distribution at the four moments marked in Fig. 1g. The sites with positive magnitude of f_i^{friction} (red colour) are the local pushing points that help the tip slip forward (to the right), while the sites with negative magnitude of f_i^{friction} (blue colour) are the local pinning points that produce lateral resistance. **e,** The histogram of $f_i^{\text{friction}}(n)$ at each of the four moments.

The histogram is obtained by dividing the range of interfacial forces into twenty bins and plotting the number of tip atoms with interaction forces within each bin. The insets provide a magnified view of the distribution in the tails. **f,** The kurtosis value for the four force distributions at the four moments. The inset shows the force distributions that were used to calculate kurtosis. $\rho(n)$ is the normalized atom number in each corresponding bin of interfacial forces.

at points iii and iv). This local force enhancement is consistent with the higher contrast of interfacial forces shown in Fig. 2a–d, which confirms the abovementioned local trap deepening. In addition, geometrical analysis also suggests that the tip–graphene interface evolved through small atomic shifts to produce more intimate atomic contact as the tip was displaced forward, providing more atomic pinning sites for lateral sliding (see Supplementary Discussion 8).

We further analysed the evolving commensurability of the interface. A less structured contact interface should have a unimodal force

distribution, similar to a Gaussian or even a uniform distribution. A very effective parameter with which to describe the deviation from a unimodal distribution (see Supplementary Discussion 9) is the kurtosis $K = \mu_4 / \sigma^4$, where μ_4 is the fourth moment about the mean value and σ is the standard deviation of the distribution (see Methods). The larger K , the more structured the distribution becomes, and the more commensurate the interface. Figure 2f show the kurtosis values for points i to iv. The kurtosis increased from point i to point iii and was almost unchanged between points iii and iv. This provides an

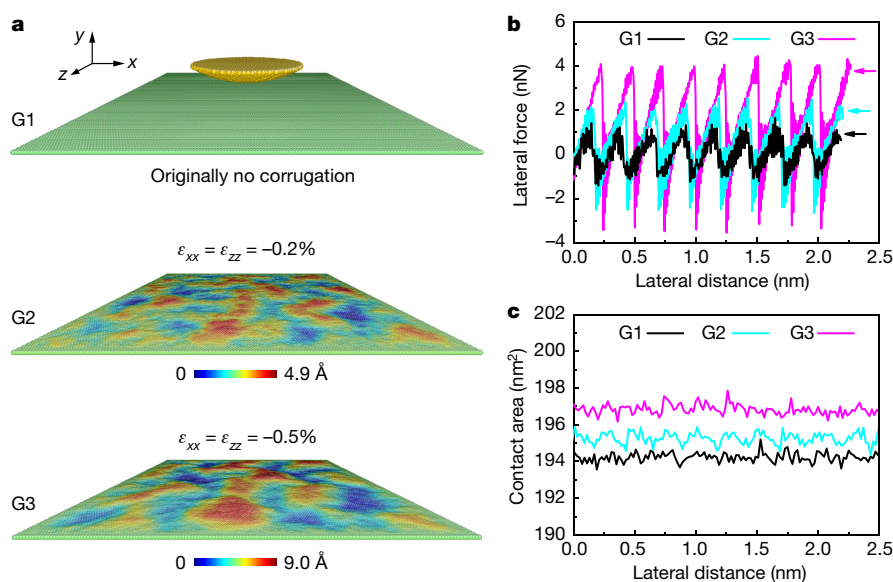


Figure 3 | Simulations of stick-slip friction on monolayer suspended graphene at 300 K. a, Simulation model. Three graphene samples with different amounts of pre-existing wrinkles are generated, named G1, G2 and G3. The colours represent the out-of-plane height amplitude.

The induced peak-to-peak height undulations of G2 and G3 are 4.9 Å and 9.0 Å, respectively. **b,** Lateral force traces showing friction sliding on the three samples. **c,** Variation of tip–graphene contact area during sliding.

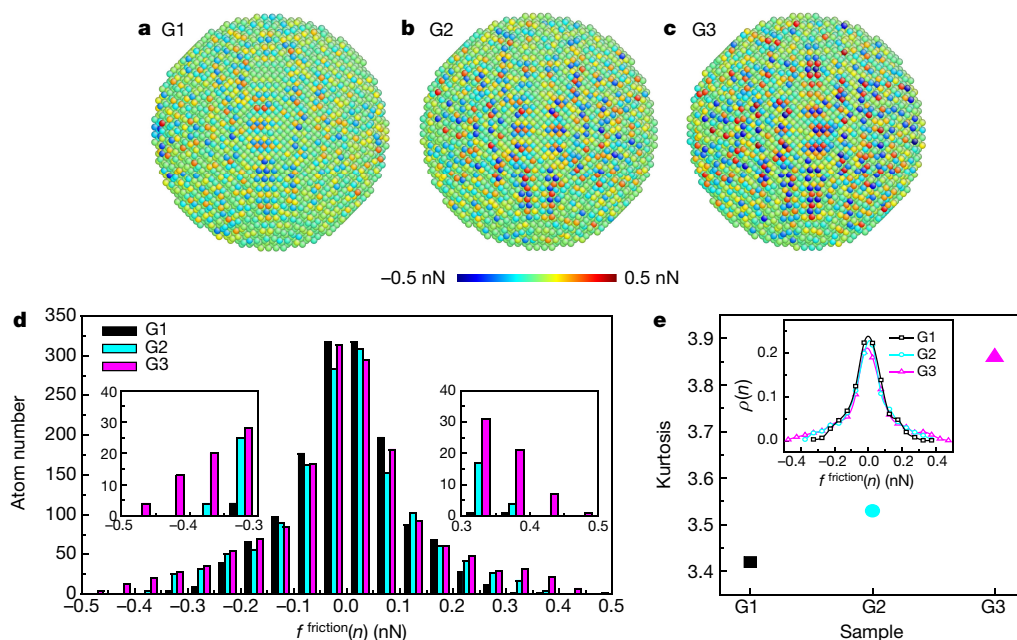


Figure 4 | Evolution of the atomic-level forces contributing to friction on monolayer suspended graphene. **a–c**, Comparison of $f_{\text{friction}}(n)$ distribution in the three samples of Fig. 3b. **d**, The histograms of $f_{\text{friction}}(n)$ for the three samples. The insets provide a magnified view of the

distribution in the tails. **e**, The kurtosis value of the distributions of $f_{\text{friction}}(n)$ for the three samples. The inset shows the force distributions that were used to calculate the kurtosis. $\rho(n)$ is the normalized atom number in each corresponding bin of interfacial forces.

indication that the tip–graphene interface became slightly more commensurate with each slip initially, and finally reached a stable value.

The simulations unambiguously show that the friction strengthening originates from two key mechanisms: enhancement of the local pinning (some individual atoms becoming more strongly pinned), and an increase in the interfacial commensurability (an increase in the number of atoms that are pinned in perfect synchrony). This is, to some extent, consistent with simulations of three-dimensional materials showing that interfacial commensurability, either through lattice matching²⁰ or through interaction-enhancing contaminant species, greatly increases static friction²¹. However, the evolution of interfacial friction and contact quality here is a result of high configurational freedom of graphene owing to its bending flexibility.

We then performed similar force analysis for multilayer ($N=2-4$) graphene systems (see Supplementary Discussion 10) and found that the variation of the interfacial pinning and overall commensurability also contributed greatly to the layer dependence of friction, explaining the large quantitative disparity between the friction enhancement and true contact area increase in Fig. 1h. However, in the multilayer cases, the top graphene layer forms a perfect AB stacking with respect to the lower graphene layer. As observed in several experiments^{2,22}, perfectly aligned stacking can strongly constrain the relative lateral movement of graphene layers. This enhanced constraint reduces the ability of graphene to adjust its configuration during the sliding process.

Since the contact interface with graphene can be affected by pre-existing wrinkles¹⁴, we speculate that regulating the degree of wrinkling²³ may offer an effective means of tuning friction. To test this, we studied a model system by sliding a tip over freely suspended graphene with various degrees of pre-existing wrinkling (Fig. 3a), formed by applying different amounts of bi-axial compressive strain (ϵ_{xx} and ϵ_{zz}). Three samples G1, G2 and G3 with 0%, -0.2%, and -0.5% pre-compressive strain were studied in the simulations. As expected, the relaxation state of graphene profoundly affects its friction (Fig. 3b). Compared to G1, relaxed samples G2 and G3 exhibited substantially enhanced static friction (up to about 300%), although the increase in true contact area was much smaller (up to about 1.5%, see Fig. 3c). The $f_{\text{friction}}(n)$ distributions at the local force peaks of these three samples (Fig. 4a–c)

and their histograms (Fig. 4d) indicate that, as pre-compressive strain increased, stronger pinning sites with higher local interfacial forces appeared. Furthermore, the kurtosis analysis shows that the interfacial force distribution increasingly deviates from a unimodal Gaussian distribution from G1 to G3 (Fig. 4e), indicating a more commensurate interface with increasing pre-compressive strain.

As confirmed by the simulations, friction on suspended graphene can be effectively tuned by pre-compression and the ‘quality’ of the interfacial contact. But unlike the supported cases, the interface reached its steady configuration immediately after the tip made contact, because of larger excess area and resultant higher configurational flexibility. Thus, no frictional strengthening was observed in these simulations, consistent with the observations of Deng *et al.*¹¹. We also performed simulations for graphene strongly bound to a flat surface. Owing to the much reduced freedom for graphene to configurationally evolve, no frictional strengthening or layer dependence occurred in this case either (see Supplementary Discussion 11), consistent with experimental measurements of strongly bound graphene^{8,24}.

In conclusion, we have identified a complementary aspect to the concept of true contact area governing friction on two-dimensional materials, related to evolving configurational relaxations that exploit out-of-plane floppiness. There is a general tendency for configurationally flexible systems to attain progressively deeper energy traps despite mechanical work being done to the system, improving the quality of contact. Imposing pre-compression on suspended graphene (to produce wrinkles) increased the total friction force several times with little change in the true contact area. This suggests a means of controlling friction of two-dimensional materials via strain engineering.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 16 February; accepted 30 September 2016.

1. Szlufarska, I., Chandross, M. & Carpick, R. W. Recent advances in single-asperity nanotribology. *J. Phys. D* **41**, 123001 (2008).
2. Dienwiebel, M. *et al.* Superlubricity of graphite. *Phys. Rev. Lett.* **92**, 126101 (2004).

3. Filippov, A. E., Dienwiebel, M., Frenken, J. W. M., Klafter, J. & Urbakh, M. Torque and twist against superlubricity. *Phys. Rev. Lett.* **100**, 046102 (2008).
4. Kim, W. K. & Falk, M. L. Atomic-scale simulations on the sliding of incommensurate surfaces: the breakdown of superlubricity. *Phys. Rev. B* **80**, 235428 (2009).
5. van Wijk, M., Dienwiebel, M., Frenken, J. & Fasolino, A. Superlubric to stick-slip sliding of incommensurate graphene flakes on graphite. *Phys. Rev. B* **88**, 235423 (2013).
6. Li, Q., Tullis, T. E., Goldsby, D. & Carpick, R. W. Frictional ageing from interfacial bonding and the origins of rate and state friction. *Nature* **480**, 233–236 (2011).
7. Liu, Y. & Szlufarska, I. Chemical origins of frictional aging. *Phys. Rev. Lett.* **109**, 186102 (2012).
8. Lee, C. *et al.* Frictional characteristics of atomically thin sheets. *Science* **328**, 76–80 (2010).
9. Chhowalla, M. & Amaratunga, G. A. Thin films of fullerene-like MoS₂ nanoparticles with ultralow friction and wear. *Nature* **407**, 164–167 (2000).
10. Choi, J. S. *et al.* Friction anisotropy-driven domain imaging on exfoliated monolayer graphene. *Science* **333**, 607–610 (2011).
11. Deng, Z. *et al.* Nanoscale interfacial friction and adhesion on supported versus suspended monolayer and multilayer graphene. *Langmuir* **29**, 235–243 (2013).
12. Cho, D.-H. *et al.* Effect of surface morphology on friction of graphene on various substrates. *Nanoscale* **5**, 3063–3069 (2013).
13. Filleter, T. *et al.* Friction and dissipation in epitaxial graphene films. *Phys. Rev. Lett.* **102**, 086102 (2009).
14. Ye, Z., Tang, C., Dong, Y. & Martini, A. Role of wrinkle height in friction variation with number of graphene layers. *J. Appl. Phys.* **112**, 116102 (2012).
15. Smolyanitsky, A., Killgore, J. & Tewary, V. Effect of elastic deformation on frictional properties of few-layer graphene. *Phys. Rev. B* **85**, 035412 (2012).
16. Ishigami, M., Chen, J. H., Cullen, W. G., Fuhrer, M. S. & Williams, E. D. Atomic structure of graphene on SiO₂. *Nano Lett.* **7**, 1643–1648 (2007).
17. Mate, C. M., McClelland, G. M., Erlandsson, R. & Chiang, S. Atomic-scale friction of a tungsten tip on a graphite surface. *Phys. Rev. Lett.* **59**, 1942–1945 (1987).
18. Bonelli, F., Manini, N., Cadelano, E. & Colombo, L. Atomistic simulations of the sliding friction of graphene flakes. *Eur. Phys. J. B* **70**, 449–459 (2009).
19. Li, Q., Lee, C., Carpick, R. W. & Hone, J. Substrate effect on thickness-dependent friction on graphene. *Phys. Stat. Sol. B* **247**, 2909–2914 (2010).
20. Luan, B. & Robbins, M. O. The breakdown of continuum models for mechanical contacts. *Nature* **435**, 929–932 (2005).
21. He, G., Müser, M. H. & Robbins, M. O. Adsorbed layers and the origin of static friction. *Science* **284**, 1650–1652 (1999).
22. Liu, Z. *et al.* Observation of microscale superlubricity in graphite. *Phys. Rev. Lett.* **108**, 205503 (2012).
23. Mohammadi, H. & Müser, M. H. Friction of wrinkles. *Phys. Rev. Lett.* **105**, 224301 (2010).
24. Spear, J. C., Custer, J. P. & Batteas, J. D. The influence of nanoscale roughness and substrate chemistry on the frictional properties of single and few layer graphene. *Nanoscale* **7**, 10021–10029 (2015).

Supplementary Information is available in the online version of the paper.

Acknowledgements S.L. and P.G. appreciate support from the Alexander von Humboldt Foundation and the Helmholtz Programme Science and Technology of Nanosystems (STN). Q.L., X.D. and J.S. appreciate support from the 973 Programs of China (grant numbers 2013CB933003, 2013CB934201, 2015CB351903 and 2012CB619402), the NSFC (grant numbers 11422218, 11272177, 11432008, 51320105014 and 51321003), the International Joint Laboratory for Micro/Nano Manufacturing and Measurement Technologies, the Tsinghua University Initiative Scientific Research Program (grant number 2014Z01007), the Thousand Young Talents Program of China and 111 project (grant number B06025). R.W.C. acknowledges support from the NSF (grant numbers CMMI-1401164 and MRSEC DMR-1120901). J.L. acknowledges support from the NSF (grant numbers MRSEC DMR-1120901, CBET-1240696, DMR-1410636 and ECCS-1610806). We also thank J. Feng for discussions.

Author Contributions Q.L., R.W.C. and J.L. conceived and designed the project. S.L. performed molecular dynamics simulations. Q.L., R.W.C. and X.Z.L. provided information about the atomic force microscope experiments. Q.L., P.G., X.D., J.S. and J.L. provided the simulation guideline. S.L., Q.L., R.W.C. and J.L. wrote the paper. All authors contributed to discussions and analyses of the results.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to Q.L. (qunyang@tsinghua.edu.cn), R.W.C. (carpick@seas.upenn.edu) or J.L. (lijun@mit.edu).

METHODS

Atomistic simulations. In the molecular dynamics simulations, we primarily studied the stick-slip friction of a tip sliding on two graphene systems: multilayer graphene supported on a rough amorphous silicon (a-Si) substrate; and suspended monolayer graphene. The covalent bonds of C–C in graphene²⁵ and Si–Si in amorphous substrate and in the crystalline Si tip²⁶ were described using the Tersoff and Stillinger–Weber potentials, respectively. A typical 6–12 Lennard–Jones potential was employed to describe van der Waals adhesive interaction between graphene and the substrate, graphene and the tip, and between the graphene layers. The Lennard–Jones parameters were chosen such that the work of adhesion (E_{ad}) or the pull-off force (f_{ad}) obtained by the molecular dynamics calculation were at the same scale as those from experimental measurements^{11,16,27–29}, as shown in Supplementary Table 1. The simulations were performed at 300 K using a Nosé–Hoover thermostat³⁰ with the LAMMPS³¹ code. The atomic configurations are displayed using AtomEye³².

For the graphene/a-Si substrate system, we first created an a-Si substrate by quenching liquid silicon from high temperature to 300 K using a cooling rate of 10^{13} K s^{−1}. The lateral dimension of the substrate was $43.4\text{ nm} \times 43.4\text{ nm}$ (x – z plane) with a thickness of approximately 3.0 nm. The atoms in the lowest 0.5 nm were held fixed. Owing to the free boundary condition along the out-of-plane (y) direction, roughness naturally formed on the free surface. A graphene sample, $38.5\text{ nm} \times 38.6\text{ nm}$ in size was then placed above the a-Si within the adhesive interaction distance, and was then allowed to relax by approaching the substrate in response to adhesion. The root-mean-square roughness for the a-Si substrate and the relaxed graphene is 1.5 Å and 1.0 Å, respectively, close to experimental values³³. After relaxation, a rigid (001)-oriented crystalline silicon tip in a bowl-shape with a radius of 16.3 nm is placed in contact with the graphene.

The rigid tip assumption was based on the following considerations. First, previous experiments⁸ have demonstrated that the layer-dependent friction on graphene was reproducible regardless of the tip material (that is, silicon, silicon nitride, or diamond). Second, using a rigid tip that is non-deformable makes it easier to trace the important yet subtle evolution of the interfacial forces.

To incorporate the compliance of atomic force microscope cantilever in experiments, we coupled harmonic springs in the x direction for pulling the tip and in the z direction for applying a fixed normal force. The stiffness of the normal springs is 30.0 N m^{-1} and 0.16 N m^{-1} , comparable to the previous experiments⁸. The entire system was relaxed for 2 ns after adding the tip and further relaxed after the normal load was imposed. The friction tests were performed by displacing the lateral spring along x direction with a constant velocity of 2 m s^{-1} with a normal load of 0.8 nN applied, and calculating the lateral force acting on the virtual atom. To enhance the damping of oscillation (primarily along the lateral direction), we artificially decreased the tip mass by a factor of ten (see Supplementary Discussion 12). Simulations were performed at 300 K. Besides the rigid tip and the fixed atoms at the boundaries, all the other atoms in graphene and a-Si substrate were subjected to the thermostat. To further justify the rigid tip assumption, we carried out another simulation using a silicon tip sliding on monolayer graphene/a-Si substrate where all tip atoms were deformable but the topmost layer of atoms were rigid. We again obtained two-stage frictional behaviour (see Supplementary Discussion 13).

In the suspended graphene system, we created an ultraflat graphene sheet with dimensions of $38.5\text{ nm} \times 38.6\text{ nm}$ in the x – z plane. The boundaries were

non-periodic. To constrain translational movement during tip sliding, the atoms near the boundaries were fixed rigidly. The tip was the same as that used in the supported system. We coupled a harmonic spring with a lateral stiffness of 30.0 N m^{-1} and pulled the silicon tip at a velocity of 2 m s^{-1} in the x direction for friction tests. To minimize the boundary effects¹⁴, a small sliding distance of 2.5 nm was used, compared to the much larger width of the graphene sheet (about 39 nm). Again, the simulations were carried out at 300 K.

Contact area. The contact area between the tip and graphene is estimated by setting a cutoff distance r_c to judge whether graphene atoms are in intimate contact with the tip atoms. Here $r_c = 4\text{ Å}$, which is slightly larger than the equilibrium distance of the Lennard–Jones potential imposed on the tip–graphene interaction. The contact area is then obtained by multiplying the total number of attached graphene atoms by the graphene atomic area of 2.77 Å^2 per atom. It should be noted that the qualitative conclusions related to contact area are not affected even when we change the cutoff distance r_c slightly. For example, the 20% increase of r_c causes an increase of contact area by only 6%.

Force analysis. To minimize the influence of thermal fluctuation on interfacial forces, the force analyses were performed on quasistatic configurations which were obtained by further relaxing the system using the conjugate gradient method. This relaxation does not alter the frictional behaviour qualitatively, but it makes the statistical results more reliable and stable.

Kurtosis. The kurtosis is calculated as $K = \mu_4/\sigma^4$, where μ_4 is the fourth moment about the mean value and σ is the standard deviation of the distribution. Specifically, $\mu_4 = \sum_n (f^{\text{friction}}(n) - \bar{f})^4 \rho(n)$ and $\sigma^2 = \sum_n (f^{\text{friction}}(n) - \bar{f})^2 \rho(n)$, where $f^{\text{friction}}(n)$ is the atomic friction force of the n th bin, $\rho(n)$ is the normalized atom number in each corresponding bin and \bar{f} is the mean value of $f^{\text{friction}}(n)$ of all bins, where $\bar{f} = \sum_n f^{\text{friction}}(n) \rho(n)$. In our case, kurtosis measures the ‘peak-enedness’ and ‘tailedness’ of the force distribution.

Data availability. The authors declare that all data supporting the findings of this study are available within the paper and its Supplementary Information.

25. Tersoff, J. Empirical interatomic potential for carbon, with applications to amorphous carbon. *Phys. Rev. Lett.* **61**, 2879–2882 (1988).
26. Stillinger, F. H. & Weber, T. A. Computer simulation of local order in condensed phases of silicon. *Phys. Rev. B* **31**, 5262–5271 (1985).
27. Zacharia, R., Ulbricht, H. & Hertel, T. Interlayer cohesive energy of graphite from thermal desorption of polyaromatic hydrocarbons. *Phys. Rev. B* **69**, 155406 (2004).
28. Zong, Z., Chen, C.-L., Dokmeci, M. R. & Wan, K.-t. Direct measurement of graphene adhesion on silicon surface by intercalation of nanoparticles. *J. Appl. Phys.* **107**, 026104 (2010).
29. Koenig, S. P., Boddeti, N. G., Dunn, M. L. & Bunch, J. S. Ultrastrong adhesion of graphene membranes. *Nat. Nanotechnol.* **6**, 543–546 (2011).
30. Nosé, S. A unified formulation of the constant temperature molecular-dynamics methods. *J. Chem. Phys.* **81**, 511–519 (1984).
31. Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **117**, 1–19 (1995).
32. Li, J. AtomEye: an efficient atomistic configuration viewer. *Model. Simul. Mater. Sci. Eng.* **11**, 173 (2003).
33. Giannazzo, F., Sonde, S., Nigro, R. L., Rimini, E. & Raineri, V. Mapping the density of scattering centers limiting the electron mean free path in graphene. *Nano Lett.* **11**, 4612–4618 (2011).

Catalytic activation of carbon–carbon bonds in cyclopentanones

Ying Xia^{1,2}, Gang Lu³, Peng Liu³ & Guangbin Dong^{1,2}

In the chemical industry, molecules of interest are based primarily on carbon skeletons. When synthesizing such molecules, the activation of carbon–carbon single bonds (C–C bonds) in simple substrates is strategically important: it offers a way of disconnecting such inert bonds, forming more active linkages (for example, between carbon and a transition metal) and eventually producing more versatile scaffolds^{1–13}. The challenge in achieving such activation is the kinetic inertness of C–C bonds and the relative weakness of newly formed carbon–metal bonds^{6,14}. The most common tactic starts with a three- or four-membered carbon-ring system^{9–13}, in which strain release provides a crucial thermodynamic driving force. However, broadly useful methods that are based on catalytic activation of unstrained C–C bonds have proven elusive, because the cleavage process is much less energetically favourable. Here we report a general approach to the catalytic activation of C–C bonds in simple cyclopentanones and some cyclohexanones. The key to our success is the combination of a rhodium pre-catalyst, an *N*-heterocyclic carbene ligand and an amino-pyridine co-catalyst. When an aryl group is present in the C3 position of cyclopentanone, the less strained C–C bond can be activated; this is followed by activation of a carbon–hydrogen bond in the aryl group, leading to efficient synthesis of functionalized α -tetralones—a common structural motif and versatile building

block in organic synthesis. Furthermore, this method can substantially enhance the efficiency of the enantioselective synthesis of some natural products of terpenoids. Density functional theory calculations reveal a mechanism involving an intriguing rhodium-bridged bicyclic intermediate.

The most common strategy for cleaving a C–C bond uses a ring system, in which the unfavourable energetics of C–C cleavage can be compensated for by the release of strain in the ring. However, although there has been success in activating C–C bonds in cyclopropane and cyclobutane derivatives (Fig. 1a)^{9–13}, the catalytic C–C activation of less strained or non-strained rings is underdeveloped. The rhodium-mediated decarbonylation of cycloalkanones has been described³; unfortunately, however, C–C activation for the less strained cyclopentanone showed low efficiency (Fig. 1b). For larger cyclic structures, an innovative chelation-based strategy has also been developed, in which ketimines prepared from the corresponding ketones and 2-amino-3-picoline allowed a directed C–C activation through a five-membered metallacycle^{8,15}. However, although medium to large cyclic ketimines can be efficiently activated in this way, the strategy is problematic for cyclopentanone- and cyclohexanone-derived ketimines (Fig. 1c)¹⁵. Our laboratory recently developed a catalytic C–C activation of isatins, a less strained ring system, to afford

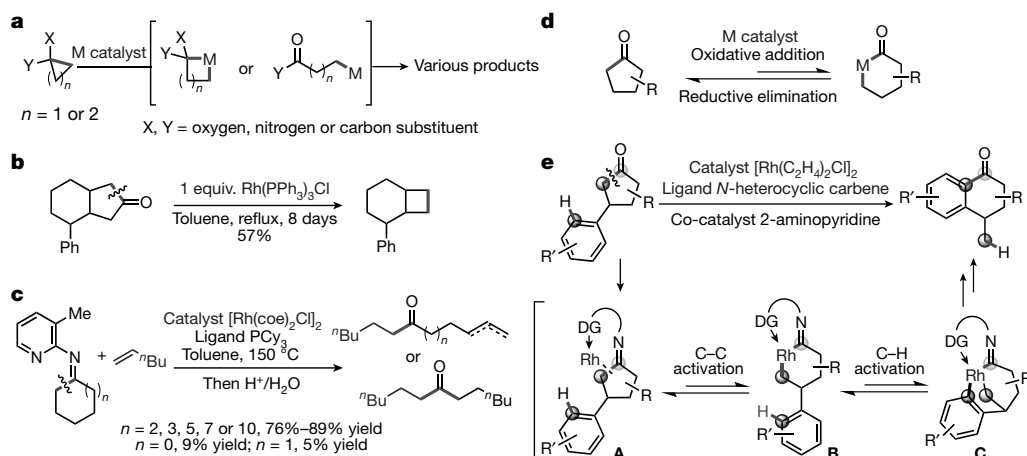


Figure 1 | Activation of C–C bonds in ring systems. **a**, Catalytic C–C activation of strained rings (for example, in cyclopropane or cyclobutanone). The unfavourable energetics of C–C activation can be compensated by the release of strain in the ring. M, transition metal. **b**, Stoichiometric rhodium-mediated C–C activation of less strained cycloketones. This reaction is less efficient than that shown in **a**. Ph, phenyl. **c**, Directed catalytic C–C activation of less strained cycloketimines. The yield is high for large cycloketimines, but low for smaller ones. Bu, butyl; coe, cyclooctene; Me, methyl; PCy₃, tricyclohexylphosphine. **d**, The challenge in terms of activating the C–C bonds of cyclopentanones is that C–C activation is reversible;

the thermodynamic driving forces do not always allow oxidative addition with a transition metal, instead favouring the reverse process (C–C reductive elimination). R, hydrocarbon side chain. **e**, Our strategy for catalytic C–C activation of cyclopentanones: merging the unfavourable C–C activation with C–H activation, to produce an overall thermodynamically favoured reaction. Specifically, when using a cyclopentanone with an aryl group in the C3 position, the imine intermediate (**A**; with a pendant directing group, DG) should allow the transient C–C activation intermediate (**B**) to undergo an intramolecular C–H activation, producing rhodacycle (**C**). Subsequent reductive elimination will lead to α -tetralone derivatives.

¹Department of Chemistry, University of Texas at Austin, Austin, Texas 78712, USA. ²Department of Chemistry, University of Chicago, Chicago, Illinois 60637, USA. ³Department of Chemistry, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA.

2-quinolinones via decarbonylation and alkyne insertion¹⁶; however, a pre-installed pyridine directing group proved essential. Hence, to our knowledge, catalytic C–C activation of normal cyclopentanones remains an unmet challenge^{17,18}.

It has been postulated that the major difficulty comes from the reversibility of the C–C activation step (Fig. 1d)^{1,6,19}. In contrast to

three- or four-membered rings, cyclopentanones lack sufficient thermodynamic driving forces to favour oxidative addition with a transition metal⁶; instead, the reverse process—C–C reductive elimination—is often preferred (Fig. 1d)^{1,19}. We hypothesized that such a problem could be solved by merging an unfavourable C–C activation with a tandem sp^2 C–H functionalization to enable an overall thermodynamically

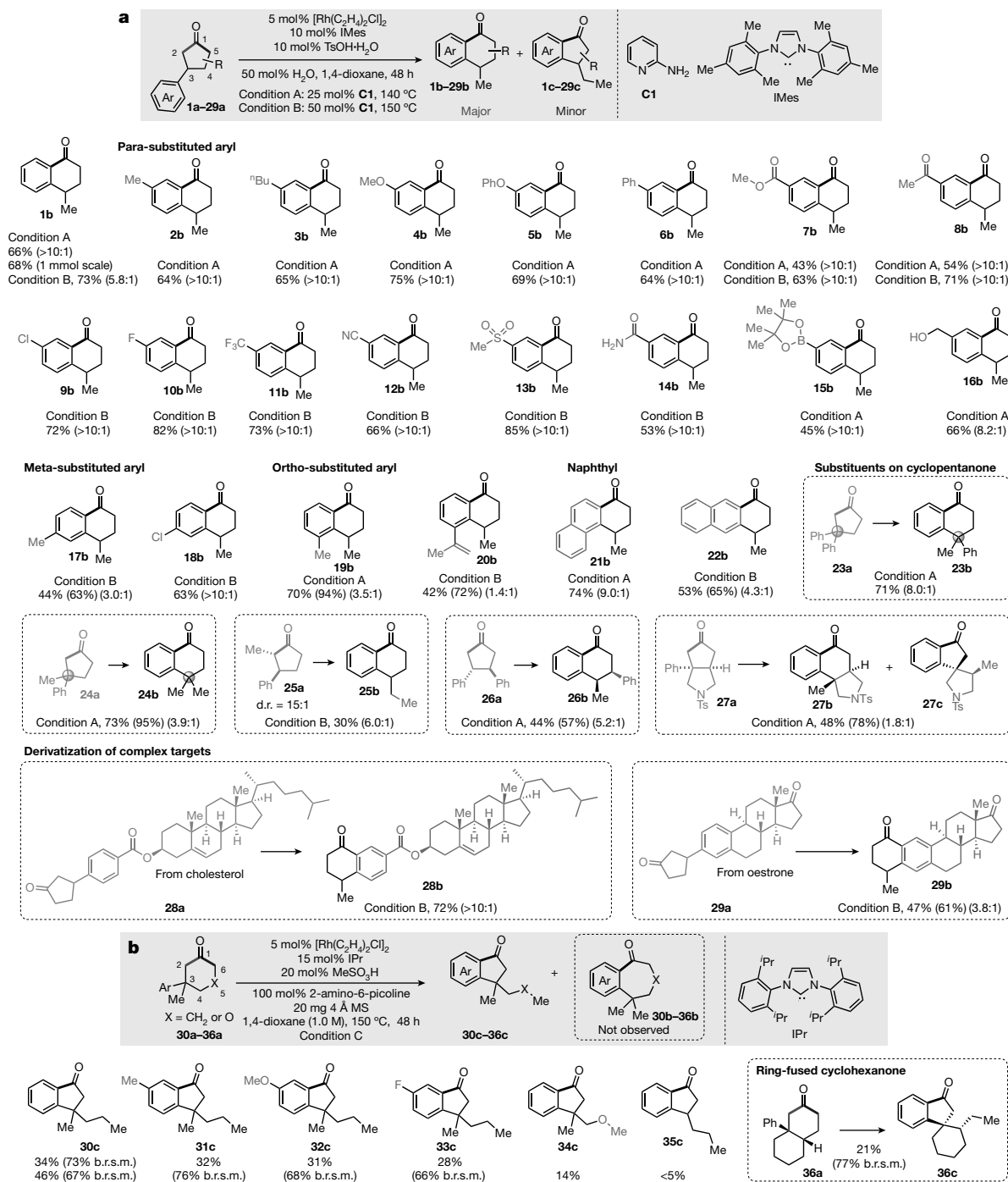


Figure 2 | Substrate scope. a, Scope of the cyclopentanones. The shaded box at the top shows the basic reaction: the substrates are 3-arylcyclopentanones (1a–29a); the major products, α -tetralones (1b–29b), come from cleavage of the more hindered C1–C2 bond, while the minor products, α -indanones (1c–29c), are generated through cleavage of the less hindered C1–C5 bond. $[Rh(C_2H_4)_2Cl]_2$ is the catalyst precursor and 2-aminopyridine (C1) is the co-catalyst; TsOH is toluene sulfonic acid. Ar, aryl group. Below the shaded box are shown the major products and their isolated yields. **b**, Scope of the cyclohexanones.

See Supplementary Information for further experimental details. The regioselective ratios (r.r.) were determined by gas chromatography-mass spectrometry or 1H nuclear magnetic resonance (NMR) of the crude products. The percentages in parentheses are the total yields or the b.r.s.m. (based on recovered starting material) yields determined by 1H NMR, using 1,1,2,2-tetrachloroethane (TCE) as the internal standard. d.r., diastereomeric ratio; IMes, 1,3-bis(2,4,6-trimethylphenyl)imidazol-2-ylidene; IPr, 1,3-bis(2,6-diisopropylphenyl)imidazol-2-ylidene; MS, molecular sieve.

favoured transformation¹³. As depicted in Fig. 1e, installing an aryl group at cyclopentanone's C3 position should allow the transient C–C activation intermediate (**B**) generated using Jun's temporary directing strategy⁸ to undergo an intramolecular ortho C–H activation, to give rhodacycle (**C**). Subsequent sp^2 – sp^2 reductive elimination would ultimately lead to α -tetralone derivatives—well known building blocks used by the chemical industry and common structural motifs found in bioactive compounds^{20–23}. The whole catalytic process would convert an aliphatic ketone to a more stable aryl ketone, accounting for the overall thermodynamic benefits of this transformation (ΔG would be about -6 kcal mol^{-1}).

To test our hypothesis, we chose 3-phenylcyclopentanone (**1a** in Fig. 2) as the model substrate. After a survey of various reaction parameters (see Supplementary Information and Supplementary Figs 1, 2), we found that using $[\text{Rh}(\text{C}_2\text{H}_4)_2\text{Cl}]_2$ as the catalyst precursor and 2-aminopyridine (**C1**) as the co-catalyst in 1,4-dioxane indeed offered the desired C–C activation products, with an 85% yield (Supplementary Table 1, entry 1). We observed two skeleton-rearranged products, in a 6.4:1 ratio. The major product, α -tetralone (**1b**), came from the cleavage of the more hindered C1–C2 bond; the minor product, α -indanone (**1c**), was generated from the cleavage of the less hindered C1–C5 bond. Further study revealed that $[\text{Rh}(\text{C}_2\text{H}_4)_2\text{Cl}]_2$ is a superior pre-catalyst, while others are much less efficient (Supplementary Table 1, entries 2–4). We investigated a series of 2-aminopyridines, and the simple 2-aminopyridine (**C1**) remained optimal; in contrast, the secondary amines (**C2** and **C3**) exhibited almost no reactivity (Supplementary Table 1, entries 5–7). In terms of ligands, the unsaturated *N*-heterocyclic carbenes IMes (1,3-bis(2,4,6-trimethylphenyl)imidazol-2-ylidene) and the bulkier IPr (1,3-bis(2,6-diisopropylphenyl)imidazol-2-ylidene) afforded comparable results (Supplementary Table 1, entry 8); in contrast, saturated *N*-heterocyclic carbenes (such as SIMes: 1,3-bis(2,4,6-trimethylphenyl)-4,5-dihydroimidazol-2-ylidene) showed substantially diminished reactivity (Supplementary Table 1, entry 9). Note that phosphine ligands, such as PCy_3 and BINAP (2,2'-bis(diphenylphosphino)-1,1'-binaphthyl), failed to promote the transformation (Supplementary Table 1, entries 10, 11). Interestingly, lowering the loading of amine co-catalyst **C1** enhanced the regioselectivity, such that the α -tetralone product was formed almost exclusively, albeit with a compromised yield (Supplementary Table 1, entry 12). Increasing the reaction temperature failed to improve the yield (Supplementary Table 1, entry 13), but addition of 50 mol% of water increased the yield to 86% without loss of regioselectivity (Supplementary Table 1, entry 14).

We then investigated the substrate scope of this C–C activation reaction (Fig. 2). Under optimized reaction conditions (condition A: 25 mol% of the co-catalyst 2-aminopyridine; temperature 140°C), standard substrate **1a** gave a somewhat lower yield when the reaction was run on a 0.2 mmol or 1.0 mmol scale. Under a slightly modified condition (condition B: 50 mol% of 2-aminopyridine; temperature 150°C), the conversion of **1a** increased and the major product (**1b**) was isolated at a yield of 73%, although the regioselective ratio (r.r.) dropped from $>10:1$ to 5.8:1. Substrates bearing an electron-neutral or electron-rich substituent at the para-position of the aryl ring afforded comparable yields (products **2b–6b**). For cyclopentanones with an electron-deficient aryl group, the reactions were much slower under condition A; however, a decent yield and excellent regioselectivity could be achieved under condition B (products **7b–14b**). A wide range of functional groups—such as ester, ketone, fluoride, aryl chloride, cyano, sulfonyl, aryl boronate, olefin and free hydroxyl groups—are well tolerated in the products (**7b–16b**, **20b**). Meta-substituted arenes are also competent substrates, in which C–H activation occurred site selectively at the less hindered position (products **17b**, **18b**). The more sterically encumbered (ortho-substituted) substrates **19a** and **20a** still exhibited high reactivity. In particular, naphthyl-substituted cyclopentanones were smoothly converted to tricyclic compounds (**21b**, **22b**); a quaternary centre at the C3 position did not affect the efficiency of

this reaction, although the regioselectivity decreased to some extent (products **23b**, **24b**). As expected, the reaction was sluggish for cyclopentanones bearing a C2 substituent (for example, **25a**); nevertheless, C–C activation still occurred predominantly at the more hindered site (to produce **25b**). Moreover, the 3,4-disubstituted and ring-fused cyclopentanones are also suitable substrates, and the relative configuration is preserved during the reaction (products **26b**, **27b**). Finally, this method allows the preparation of unusual derivatives of complex bioactive molecules (**28b** and **29b**). For example, a fused ring can be conveniently installed on oestrone (to produce **29b**).

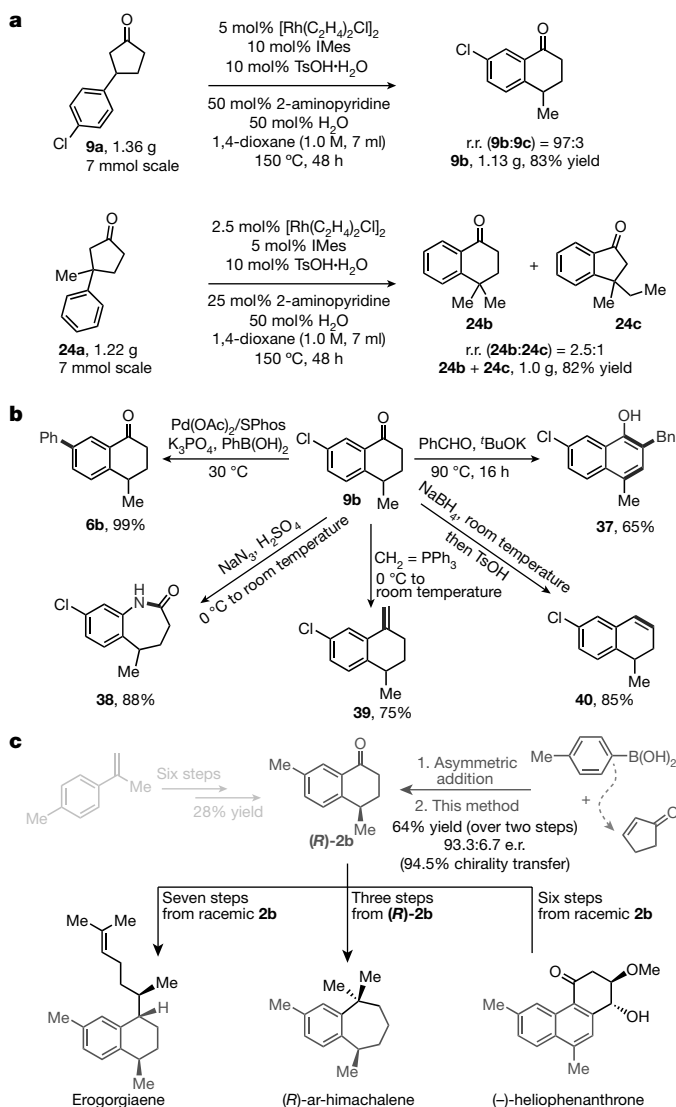


Figure 3 | Gram-scale synthesis and synthetic applications. **a**, Gram-scale reactions. Compared with the small-scale reaction, using substrate **9a** in a gram-scale reaction produced a higher percentage yield; similarly, when using substrate **24a**, less catalyst was needed. **b**, Standard protocols can be used to add several new functional groups to our C–C activation product α -tetralone **9b**. Bn, benzyl group. **c**, Applications in asymmetric total syntheses of terpenoids. It is known that α -tetralone (**2b**) is an intermediate for accessing erogorgiaene^{20,21}, (*R*)-ar-himachalene²² and (–)-heliophenanthrone²³ in three to seven steps (shown at the bottom). However, enantioselective preparation of α -tetralones with a C4 stereocentre is non-trivial and generally requires many steps²² (top left, pale grey). In our technique (top right, dark grey), we synthesized optically enriched 3-arylcyclopentanones in a single step through asymmetric conjugate addition of cyclopentenone and arylboronic acids²⁴; then, using our C–C activation approach, we isolated α -tetralone (**(R)-2b**) in 64% yield over two steps, with 94.5% chirality transfer. e.r., enantiomeric ratio.

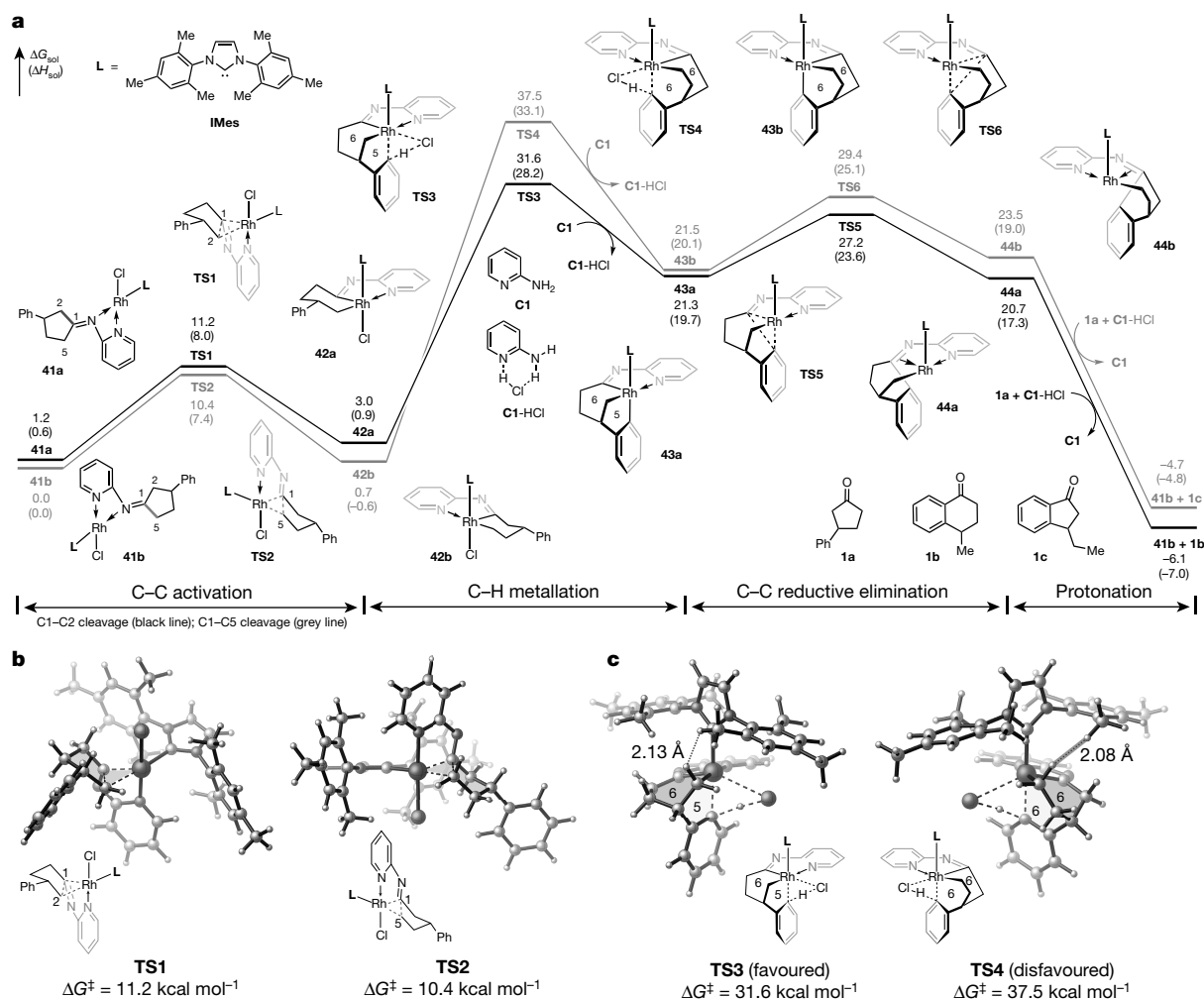


Figure 4 | DFT-computed pathways for the activation of C–C bonds in cyclopentanones. **a**, The computed reaction energy profile of the reaction with cyclopentanone **1a**. **L**, ligand; **TS**, transition state; $\Delta G_{\text{sol}}^{\ddagger}$, Gibbs free energy with respect to **41b** (given in kcal mol⁻¹); $\Delta H_{\text{sol}}^{\ddagger}$, enthalpy with respect to **41b** (given in kcal mol⁻¹). **b**, Transition states during

C–C bond activation. ΔG^{\ddagger} , Gibbs free energy of activation. **c**, Transition states during C–H bond cleavage. Energies are computed at the M06/SDD-6-311+G(d,p)/SMD (1,4-dioxane) level of theory, with geometries optimized at the B3LYP/LANL2DZ-6-31G(d) level (see Supplementary Information for more details and references).

Compared with cyclopentanones, cyclohexanones are generally less strained (see the computed ring strains in Supplementary Information), and thus are more challenging substrates for C–C activation¹⁵. Indeed, under our standard conditions (A or B), no C–C activation or rearrangement product was observed when using 3-arylcyclohexanones (for example, **30a** or **35a**) as the substrate. However, when using the bulkier IPr as the ligand and 2-amino-6-picoline as the co-catalyst (Fig. 2b, condition C), the C–C activation product **30c** was isolated in 34%–46% yield (67%–73% yield on the basis of recovered starting material). For example, the C1–C6 bond of cyclohexanone **30a** was exclusively cleaved to furnish the product α -indanone, giving the opposite regioselectivity from cyclopentanones. In addition, cyclohexanone substrates containing electron-rich or electron-deficient aryl groups show comparable reactivity (products **31c**–**33c**). Moreover, heterocyclic substrates, such as **34a**, also underwent the C–C activation/rearrangement to produce alkoxy-substituted α -indanone (**34c**). The quaternary centre in the cyclohexanones appears to facilitate the C–C/C–H activation process, as 3-phenylcyclohexanone (**35a**) only afforded a trace amount of the product (**35c**). The 6,6-fused cyclohexanone **36a** can also undergo this transformation to afford a spirocyclic α -indanone (**36c**).

To demonstrate the practicability of this transformation, we carried out gram-scale reactions (Fig. 3a), which offered either a higher yield compared with the small-scale reaction (when using substrate **9a**; see

above and Fig. 2) or a lower catalyst loading (when using substrate **24a**). To show the versatile nature of α -tetralones in organic synthesis, we introduced several new functional moieties to C–C activation product **9b** through standard protocols (Fig. 3b). The synthetic potential of this method can be further demonstrated in the asymmetric syntheses of natural products (Fig. 3c). For example, α -tetralone (**2b**) is a known intermediate for accessing erogorgiaene^{20,21}, (*R*)-ar-himachalene²² and (–)-heliophenanthrone²³ in three to seven steps; however, enantioselective preparation of α -tetralones with a C4 stereocentre is non-trivial and generally requires many steps²². Given that optically enriched 3-arylcyclopentanones can be synthesized in a single step from asymmetric conjugate addition between cyclopentenone and arylboronic acids²⁴, using the C–C activation approach we isolated α -tetralone (**(R)**-**2b**) in 64% yield over two steps, with 94.5% chirality transfer. The slight decrease in enantioselectivity was probably caused by a reversible β -H elimination process with the C–C activation intermediate (see rhodacycle **B** above; Fig. 1e). With an efficient route to (*R*)-**2b**, the synthesis of the aforementioned terpenoids is markedly streamlined.

To better understand the origin of the regioselectivity and the mechanism of the C–H metallation step, we carried out density functional theory (DFT) calculations. The computed free energy profile of the key steps in the catalytic cycle is shown in Fig. 4a. Through the

installation of a temporary directing group, the resulting ketimines (both *E* and *Z* isomers) can initially bind to the rhodium in a bidentate fashion to form **41a** and **41b**. The pyridyl group remains bound to the metal and promotes C–C oxidative addition (via transition states **TS1** and **TS2**, Fig. 4b). While cleavage of the sterically more hindered C1–C2 bond in **TS1** is slightly less favoured, activations of the C1–C2 and C1–C5 bonds are both reversible, with a relatively low activation barrier. Indeed, the regioselectivity is determined in the subsequent C–H metallation step (via **TS3** and **TS4**) and C–C reductive elimination step (via **TS5** and **TS6**)—both steps requiring much higher activation energies than that of the C–C cleavage. The pathway with rhodacycle **42a** (derived from activation of the C1–C2 bond) has lower barriers in both C–H metallation (**TS3**) and C–C reductive elimination (**TS5**) than the corresponding reaction with rhodacycle **42b** (derived from activation of the C1–C5 bond).

The substantial destabilization of the disfavoured transition states (**TS4** and **TS6**) is caused mainly by the steric repulsions between the C5 methylene group in the forming six-membered rhodacycle and the ortho methyl group on the IMes ligand. This unfavourable interaction in **TS4** is evidenced by the short hydrogen–hydrogen distance of 2.08 Å (Fig. 4c). In contrast, these steric repulsions are diminished in **TS3** and **TS5**, in which the methylene group is positioned further away from the IMes ligand owing to the shorter tether in the forming five-membered metallacycle. The strong kinetic preference for the pathway involving the 5,6-bridged rhodacycle **43a** explains the high level of regioselectivity for the C1–C2-activation product. The subsequent protonation (Supplementary Fig. 12) requires lower barriers and provides thermodynamic driving forces to finally form α -tetralone (**1b**) upon hydrolysis. Regarding the C–H metallation step, we found that a chloride-mediated metallation-deprotonation pathway²⁵ was more favourable; in contrast, concerted 1,4-rhodium migration^{26–30} through σ -bond metathesis can be ruled out owing to the much higher activation energy required.

This computationally proposed mechanism is also consistent with observed kinetic isotope effects and with isotope labelling experiments (see Supplementary Information). First, the small primary kinetic isotope effect suggests that the C–H metallation and C–C reductive elimination are both turnover limiting. Second, the incomplete deuterium transfer seen in a reaction with deuterated substrate provides further evidence against the concerted 1,4-rhodium migration mechanism.

Received 27 December 2015; accepted 6 September 2016.

Published online 2 November 2016.

- Gozin, M., Weisman, A., Ben-David, Y. & Milstein, D. Activation of a carbon–carbon bond in solution by transition-metal-insertion. *Nature* **364**, 699–701 (1993).
- Jones, W. D. The fall of the C–C bond. *Nature* **364**, 676–677 (1993).
- Murakami, M., Amii, H. & Ito, Y. Selective activation of carbon–carbon bonds next to a carbonyl group. *Nature* **370**, 540–541 (1994).
- Rybtchinski, B. & Milstein, D. Metal insertion into C–C bonds in solution. *Angew. Chem. Int. Ed.* **38**, 870–883 (1999).
- Murakami, M. & Ito, Y. Cleavage of carbon–carbon single bonds by transition metals. *Top. Organomet. Chem.* **3**, 97–129 (1999).
- Jun, C.-H. Transition metal-catalyzed carbon–carbon bond activation. *Chem. Soc. Rev.* **33**, 610–618 (2004).
- Miura, M. & Satoh, T. Catalytic processes involving β -carbon elimination. *Top. Organomet. Chem.* **14**, 1–20 (2005).
- Park, Y. J., Park, J.-W. & Jun, C.-H. Metal–organic cooperative catalysis in C–H and C–C bond activation and its concurrent recovery. *Acc. Chem. Res.* **41**, 222–234 (2008).
- Dong, G. (ed.) *C–C bond activation* Vol. 346 of *Topics in Current Chemistry* (Springer, 2014).
- Souillart, L. & Cramer, N. Catalytic C–C bond activations via oxidative addition to transition metals. *Chem. Rev.* **115**, 9410–9464 (2015).
- Seiser, T., Saget, T., Tran, D. N. & Cramer, N. Cyclobutanes in catalysis. *Angew. Chem. Int. Ed.* **50**, 7740–7752 (2011).

- Mack, D. J. & Njardarson, J. T. Recent advances in the metal-catalyzed ring expansions of three- and four-membered rings. *ACS Catal.* **3**, 272–286 (2013).
- Masarwa, A. *et al.* Merging allylic carbon–hydrogen and selective carbon–carbon bond activation. *Nature* **505**, 199–203 (2014).
- Halpern, J. Determination and significance of transition-metal-alkyl bond dissociation energies. *Acc. Chem. Res.* **15**, 238–244 (1982).
- Jun, C.-H., Lee, H. & Lim, S.-G. The C–C bond activation and skeletal rearrangement of cycloalkanone imine by Rh(I) catalysts. *J. Am. Chem. Soc.* **123**, 751–752 (2001).
- Zeng, R. & Dong, G. Rh-catalyzed decarbonylative coupling with alkynes via C–C activation of isatins. *J. Am. Chem. Soc.* **137**, 1408–1411 (2015).
- Mukai, C., Ohta, Y., Oura, Y., Kawaguchi, Y. & Inagaki, F. Csp³–Csp³ and Csp³–H bond activation of 1,1-disubstituted cyclopentane. *J. Am. Chem. Soc.* **134**, 19580–19583 (2012).
- Li, G., Arisawa, M. & Yamaguchi, M. Rhodium-catalyzed synthesis of unsymmetrical di(aryl/heteroaryl)methanes using aryl/heteroarylmethyl ketones via CO–C bond cleavage. *Chem. Commun.* **50**, 4328–4330 (2014).
- Albrecht, M., Gossage, R. A., Spek, A. L. & van Koten, G. Metal-mediated C–C bond making and breaking: first direct evidence for a reversible migration of a benzyl group along a metal–carbon bond. *J. Am. Chem. Soc.* **121**, 11898–11899 (1999).
- Davies, H. M. L. & Walji, A. M. Direct synthesis of (+)-erogorgiaene through a kinetic enantiodifferentiating step. *Angew. Chem. Int. Ed.* **44**, 1733–1735 (2005).
- Elford, T. G., Nave, S., Sonawane, R. P. & Aggarwal, V. K. Total synthesis of (+)-erogorgiaene using lithiation/borylation methodology, and stereoselective synthesis of each of its diastereoisomers. *J. Am. Chem. Soc.* **133**, 16798–16801 (2011).
- Chavan, S. P. & Khatod, H. S. Enantioselective synthesis of the essential oil and pheromonal component ar-himachalene by a chiral pool and chirality induction approach. *Tetrahedron Asymmetry* **23**, 1410–1415 (2012).
- Mukherjee, P. & Sarkar, T. K. Heteroatom-directed Wacker oxidations. A protection-free synthesis of (–)-heliophenanthrone. *Org. Biomol. Chem.* **10**, 3060–3065 (2012).
- Hayashi, T. & Yamasaki, K. Rhodium-catalyzed asymmetric 1,4-addition and its related asymmetric reactions. *Chem. Rev.* **103**, 2829–2844 (2003).
- Lapointe, D. & Fagnou, K. Overview of the mechanistic work on the concerted metallation deprotonation pathway. *Chem. Lett.* **39**, 1118–1126 (2010).
- Matsuda, T., Shigeno, M. & Murakami, M. Asymmetric synthesis of 3,4-dihydrocoumarins by rhodium-catalyzed reaction of 3-(2-hydroxyphenyl) cyclobutenones. *J. Am. Chem. Soc.* **129**, 12086–12087 (2007).
- Seiser, T., Roth, O. A. & Cramer, N. Enantioselective synthesis of indanols from *tert*-cyclobutanols using a rhodium-catalyzed C–C/C–H activation sequence. *Angew. Chem. Int. Ed.* **48**, 6320–6323 (2009).
- Shigeno, M., Yamamoto, T. & Murakami, M. Stereoselective restructuring of 3-arylcyclobutanols into 1-indanols by sequential breaking and formation of carbon–carbon bonds. *Chem. Eur. J.* **15**, 12929–12931 (2009).
- Matsuda, T., Suda, Y. & Takahashi, A. Double 1,4-rhodium migration cascade in rhodium-catalysed arylative ring-opening/spirocyclisation of (3-arylcyclobutylidene)acetates. *Chem. Commun.* **48**, 2988–2990 (2012).
- Matsuda, T. & Yuihara, I. A rhodium(I)-catalysed formal intramolecular C–C/C–H bond metathesis. *Chem. Commun.* **51**, 7393–7396 (2015).

Supplementary Information is available in the online version of the paper.

Acknowledgements This project was supported by the Cancer Prevention Research Institute of Texas (grant R1118), the National Institute of General Medical Science (grant R01GM109054) and the Welch Foundation (grant F1781). G.D. is a Searle Scholar and Sloan fellow. Y.X. acknowledges the International Postdoctoral Exchange Fellowship Program 2015 from the Office of China Postdoctoral Council (OCP, document 38, 2015). We thank Johnson Matthey for a donation of Rh salts, and Chiral Technologies for donation of chiral high-performance liquid-chromatography columns. We are grateful to Y. Xu for providing 1,4-dioxane, F. Mo for providing some 3-aryl cyclopentanones, and H. Lim for checking the experimental procedures. DFT calculations were performed using supercomputer resources at the Center for Simulation and Modeling at the University of Pittsburgh, and the Extreme Science and Engineering Discovery Environment supported by the National Science Foundation.

Author Contributions Y.X. and G.D. conceived and designed the experiments. Y.X. performed the experiments. G.L. performed the DFT calculations. Y.X., G.L., P.L. and G.D. co-wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.L. (pengliu@pitt.edu) and G.D. (gbdong@uchicago.edu).

Reviewer Information *Nature* thanks J. Harvey, M. Lautens and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Magnetic reversals from planetary dynamo waves

Andrey Sheyko¹, Christopher C. Finlay² & Andrew Jackson¹

A striking feature of many natural dynamos is their ability to undergo polarity reversals^{1,2}. The best documented example is Earth's magnetic field, which has reversed hundreds of times during its history^{3,4}. The origin of geomagnetic polarity reversals lies in a magnetohydrodynamic process that takes place in Earth's core, but the precise mechanism is debated⁵. The majority of numerical geodynamo simulations that exhibit reversals operate in a regime in which the viscosity of the fluid remains important, and in which the dynamo mechanism primarily involves stretching and twisting of field lines by columnar convection⁶. Here we present an example of another class of reversing-geodynamo model, which operates in a regime of comparatively low viscosity and high magnetic diffusivity. This class does not fit into the paradigm of reversal regimes that are dictated by the value of the local Rossby number (the ratio of advection to Coriolis force)^{7,8}. Instead, stretching of the magnetic field by a strong shear in the east–west flow near the imaginary cylinder just touching the inner core and parallel to the axis of rotation is crucial to the reversal mechanism in our models, which involves a process akin to kinematic dynamo waves^{9,10}. Because our results are relevant in a regime of low viscosity and high magnetic diffusivity, and with geophysically appropriate boundary conditions, this form of dynamo wave may also be involved in geomagnetic reversals.

Motion of liquid metal in Earth's core is the source of the geomagnetic field and its time variations, including dramatic polarity reversals wherein the large-scale field switches to point in the opposite direction. The past twenty years have seen considerable advances in our ability to model this process, with the advent of 3D numerical geodynamo models capable of solving the equations of conservation of momentum, magnetic induction and buoyancy transport in a self-consistent manner^{11,12}. Some such models show polarity reversals^{13–15}, which proceed by the production and expulsion of inverse magnetic flux by intermittent convective plumes, and subsequent transport by meridional flow^{6,16}. By necessity, such simulations operate in a regime in which viscous effects remain important. However, the values of the two dimensionless parameters that characterize these simulations—the Ekman number $Ek = \nu / (2\Omega d^2)$ and the magnetic Prandtl number $Pr_m = \nu / \eta$, where ν is the kinematic viscosity, Ω is the angular rotation rate, d is the thickness of the spherical shell of convecting fluid and η is the magnetic diffusivity—differ greatly from the values estimated for Earth's core, and so it is possible that other processes operate under such conditions¹⁷. Figure 1 summarizes Ek and Pr_m for a selection of published models that exhibit polarity reversals (see Extended Data Table 1).

In an effort to push towards the low-viscosity regime, some researchers have experimented with stress-free boundary conditions in geodynamo models, and discovered a wide variety of quadrupolar and hemispheric dynamos, some of which undergo prominent field oscillations, excursions and reversals^{10,18} via a dynamo-wave mechanism⁹. Dynamo waves have been extensively studied in both 2D and 3D using prescribed flows^{2,19,20}; they involve cyclic variations in magnetic induction due to the combined action of differentially rotating (shear) flow and helical flow. Up to now, the relevance of such solutions to Earth has been uncertain because they have been found only in kinematic models,

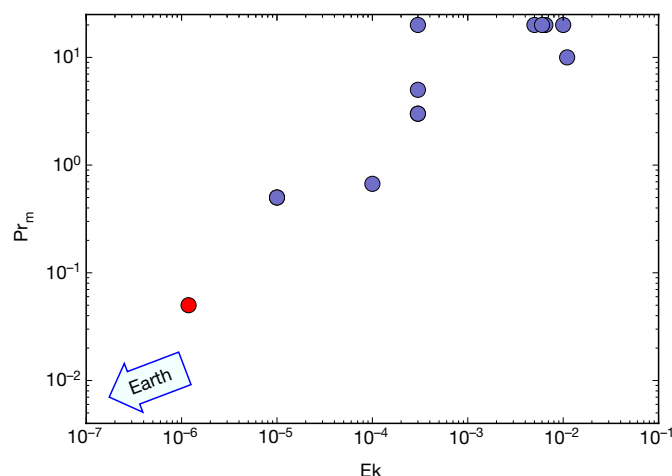


Figure 1 | Control parameters of a selection of numerical geodynamo models that exhibit polarity reversals. The blue dots are the parameters for published reversing-geodynamo models (see Extended Data Table 1). The red dot shows the parameters for the reversing dynamo S6 discussed here. $Ek = \nu / (2\Omega d^2)$ is the Ekman number and $Pr_m = \nu / \eta$ is the magnetic Prandtl number, where ν is the kinematic viscosity, Ω is the angular rotation rate, d is the thickness of the spherical shell of convecting fluid and η is the magnetic diffusivity. For case S6, $Ek = 1.2 \times 10^{-6}$ and $Pr_m = 0.05$. The Earth lies off to the bottom left in this diagram (as indicated by the arrow), with estimated values of $Ek = \mathcal{O}(10^{-15})$ and $Pr_m = \mathcal{O}(10^{-5})$.

or in dynamic models with unusual boundary conditions or forcing modes. Here, we demonstrate that dynamo waves can also produce polarity reversals in a standard geodynamo model set-up, provided that Ek and Pr_m are sufficiently small and that the rate of convective driving is sufficiently large.

We adopt a numerical approach to solve the equations governing convection-driven magnetohydrodynamics in a rapidly rotating, electrically conducting, spherical shell (see Methods). The calculations reported here were started from an earlier run with parameters and set-up as close as possible to a previously reported dynamo simulation^{21,22}, which uses among the smallest published values for Ek (1.2×10^{-6}) and Pr_m (0.2). In a calculation hereafter referred to as 'S6', the convective driving (as measured by the Rayleigh number Ra ; see Equation (2) in Methods) was then increased by a factor of 30 and Pr_m decreased to 0.05, in an effort to break the strong equatorial anti-symmetry of the generated field and to enhance its time variability. Figure 2a shows that, following the change in the control parameters, there is a gradual decrease in the magnitude of the magnetic energy in dynamo S6. Then, after about 1.1 magnetic diffusion times with a steady, dipole-dominated field, nearly periodic reversals of the dipole field set in (Fig. 2b). The period of the reversals is approximately 0.18 magnetic diffusion times and they persist for the 6 magnetic diffusion times that we have been able to calculate so far. We repeated the calculations with increased model resolution, in both angular and radial directions, and continue to find the same reversing behaviour.

¹Institute of Geophysics, ETH Zurich, 8092 Zurich, Switzerland. ²Division of Geomagnetism, DTU Space, Technical University of Denmark, 2800 Kongens Lyngby, Denmark.

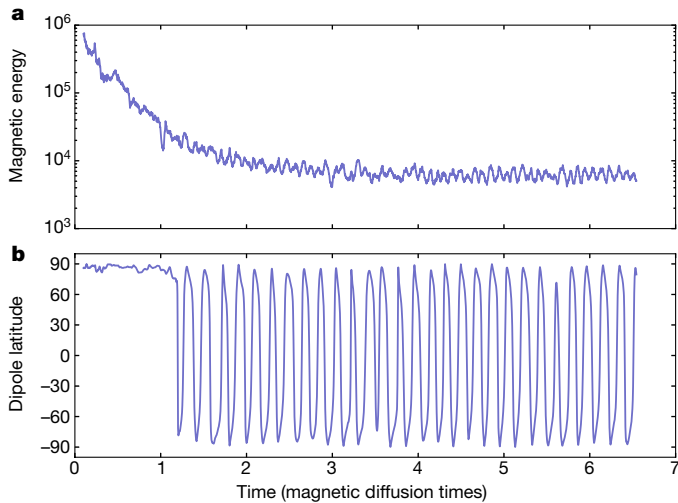


Figure 2 | Magnetic field time-dependence for dynamo S6. **a**, Time dependence of the magnetic energy, integrated over the volume of the dynamo region, for dynamo S6. The simulation was started from a previous dipole-dominated run similar to the dynamo of ref. 21, then the Rayleigh number Ra was increased by a factor of 30 and the magnetic Prandtl number Pr_m decreased by a factor of 4. **b**, The resultant time dependence of the dipole tilt, determined from the first three Gauss coefficients $\{g_1^0, g_1^1, h_1^1\}$ of the magnetic field at the outer boundary. Polarity reversals set in around 1.1 magnetic diffusion times.

Figure 3 presents the detailed structure and time-dependence of the reversing field in dynamo S6. The radial magnetic field B_r at the outer boundary of the dynamo is primarily axisymmetric, and consists of bands with alternating polarity that appear at low latitude and subsequently move polewards (Fig. 3b, c and Supplementary Video 1). The northern hemisphere field is stronger and migrates polewards

more slowly (see the lines in Fig. 3b, which illustrate that the speed dz/dt , where $z = \sin(\lambda)$ and λ is the latitude, is 50% faster in the southern hemisphere). The reversing field at the outer boundary is linked to deep structures that are evident in meridional sections of the longitudinally averaged azimuthal field \bar{B}_ϕ (Fig. 3d and Supplementary Videos 2, 3). In the equatorial plane, the sign of \bar{B}_ϕ alternates, and the field subsequently migrates polewards in each hemisphere, moving parallel to the rotation axis.

In Fig. 4, we examine the underlying flow that generates this distinctive pattern of field reversals. Figure 4a shows the time-averaged kinetic energy spectrum as a function of spherical harmonic degree. The kinetic energy in the reversing regime is much larger than the magnetic energy, and it decays more slowly with spherical harmonic degree. It is nonetheless dominated by degrees below 20, and has a notably large zonal component. Snapshots and animations of the flow (Supplementary Videos 4, 5) reveal columnar flow patterns with strong alignment parallel to the rotation axis, typical of low-Ek rotating convection. The time and longitudinally averaged flow in S6 is presented in Fig. 4b, c, which shows meridional sections of the angular velocity (differential rotation) and radial flow. There is a strong retrograde jet close to the tangent cylinder; this produces substantial shear perpendicular to the rotation axis and is associated with weaker, but persistent, radial flow. There is a surprising difference between the angular velocities in the northern and southern hemispheres: prograde zonal flow at low latitudes in the southern hemisphere is absent in the northern hemisphere. This may be linked to the southern hemisphere being colder at the outer boundary (see Supplementary Video 6). Flow asymmetry between the northern and southern hemispheres, and a substantial quadrupolar component in the generated field, are conditions known to favour reversing behaviour in dynamos^{23,24}. Neither the kinetic energy nor the structure of the flows varies much during the reversals.

To further investigate the nature of the reversal process, we carried out a series of additional calculations, in which we varied the control

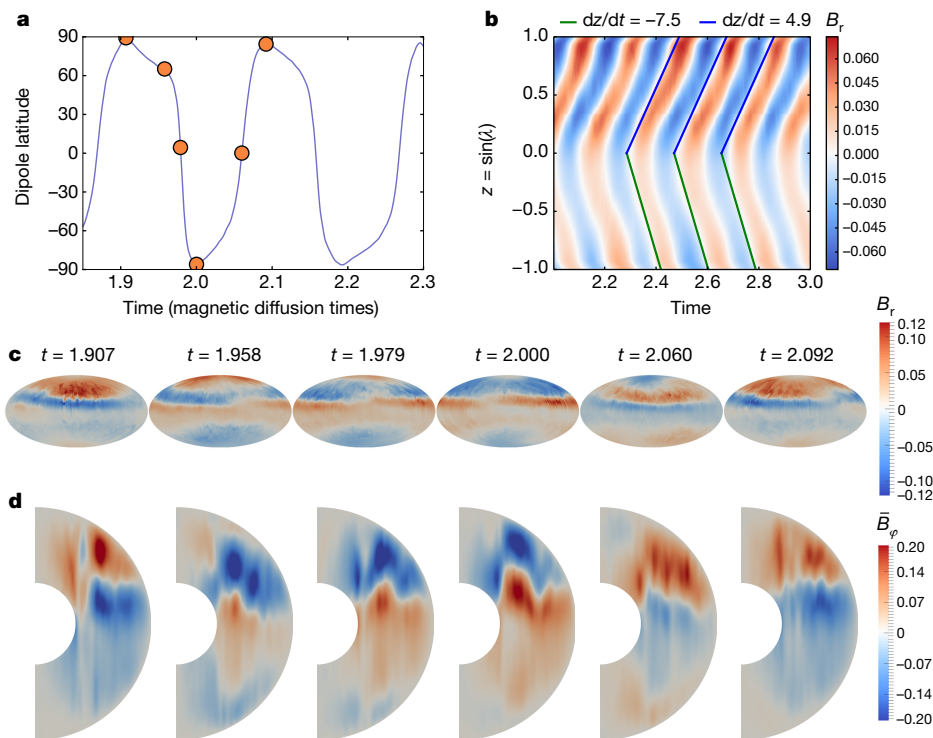


Figure 3 | Magnetic field polarity reversal process. **a**, A close up of time dependence of the dipole tilt angle from dynamo S6 (Fig. 2b); orange dots mark the time instances shown in **c** and **d**. **b**, A butterfly plot of the longitudinally averaged radial magnetic field B_r at the outer boundary of the dynamo region as a function of $z = \sin(\lambda)$, where λ is the latitude,

and time **c**, Sequence showing B_r at the outer boundary of the liquid metal region, in a Hammer projection, for the times marked by the orange dots in **a**. **d**, Sequence of meridional sections showing the longitudinally averaged azimuthal field \bar{B}_ϕ within the liquid metal region.

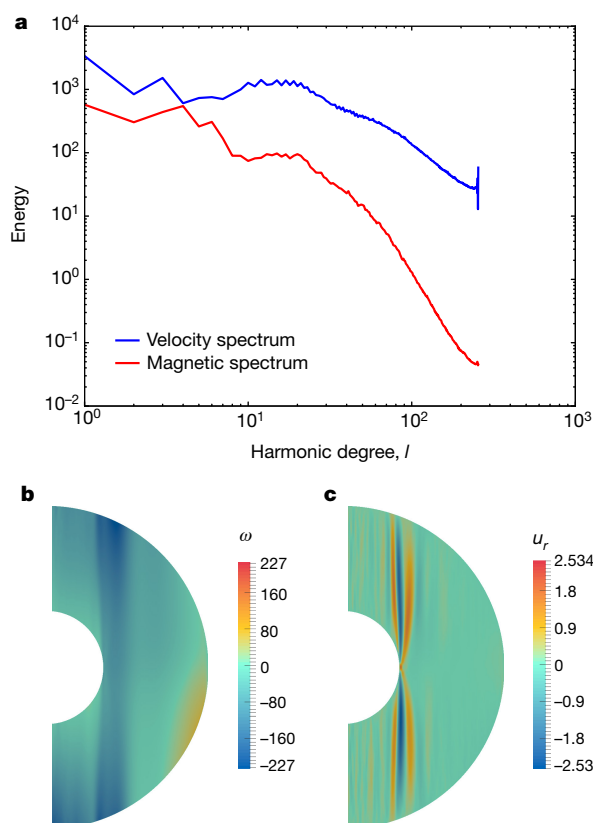


Figure 4 | Time-averaged flow in dynamo S6. **a**, Volume-averaged spectra of the kinetic energy (blue) compared to the magnetic energy (red), time-averaged over the reversing interval. **b**, A meridional section showing the time and longitudinally averaged angular velocity ω , with blues denoting retrograde velocities and reds denoting prograde velocities. **c**, Similar meridional section for the radial flow (u_r , radial velocity); reds show radially outward flow.

parameters and set-up of the dynamo (see Extended Data Table 2). Upon increasing Ra , which enhances the magnetic Reynolds number Re_m and the amplitude of the zonal flow, we continue to find similar reversals, but with shorter periods and increasing time variability (runs S6.04–S6.08). Decreasing Pr_m much further is difficult without losing the dynamo, but small decreases slightly shorten the reversal period (runs S7.02, S7.03). Replacing the electrically conducting inner core with an insulating inner core makes little difference (the reversal period becomes very slightly shorter; see run S6_InsIC), presumably because the reversals result primarily from processes outside the tangent cylinder. We also carried out one calculation with the Lorentz force switched off (case S6_LorOff). Very similar reversals were then obtained, but with a shorter period of only 0.07 magnetic diffusion times. The helicity of the flow did not change much upon switching off the Lorentz force, but the energy in the zonal part of the toroidal flow increased by a factor of five. A calculation using a revised model ‘S6_{ε0}’, in which the internal heating is set to zero (and all other parameters remain the same as in S6; see Extended Data Table 2), has established that in this case the dynamo continues to reverse in a similar manner, with its period differing by less than 5% from that of S6 (see Methods).

We next consider whether these reversals can be the result of a dynamo-wave process. A first clue is that Parker’s theory^{2,9} predicts poleward propagation of dynamo waves along cylinders parallel to the rotation axis, in the case that the magnitude of the (westward) angular velocity decreases with distance away from the tangent cylinder and helicity is negative in the northern hemisphere. This is in agreement with our results. In addition, the period of dynamo waves can be estimated by^{10,18}

$$T_{\text{Parker}} = 2\pi\gamma[H(\overline{E_{\text{tor}}})^{1/2}]^{-1/2} \quad (1)$$

where T_{Parker} is the predicted reversal period, γ is a constant, $H = \left| \int_V (\mathbf{u} - \bar{\mathbf{u}}) \cdot \nabla \times (\mathbf{u} - \bar{\mathbf{u}}) dV \right|$ is the modulus of the volume-integrated kinetic helicity of the non-axisymmetric flow that we evaluate in the northern hemisphere, and $\overline{E_{\text{tor}}}$ is the zonal toroidal flow energy. Both quantities were time-averaged over the reversing interval. Upon comparing the predictions of this theory with the reversal periods found in our numerical calculations (Extended Data Table 2) we find a reasonable agreement (Extended Data Fig. 1), especially regarding the decrease in reversal period as the energy of toroidal zonal flow increases.

Dynamo S6 thus demonstrates that field oscillations and reversals can result from dynamo waves produced by rapidly rotating convection in a spherical shell, even in the presence of no-slip boundary conditions, provided that Ek and Pr_m are sufficiently small. Dynamo waves are apparently a very robust phenomenon, relevant to dynamos across a range of planetary^{25–27} and astrophysical²⁸ scenarios. A common aspect of such dynamos is that they possess a relatively large toroidal magnetic field, obtained by stretching due to differential rotation. Strong azimuthal fields just outside the tangent cylinder, which are related to an omega effect taking place at this location, are a characteristic feature of the reversing dynamos reported here.

Our more Earth-like parameter regime brings further surprises. The conventional wisdom is that the local Rossby number Ro_l (see Methods) is important for delineating two regimes of behaviour⁷: for $Ro_l < 0.1$ the dynamos are generally dipolar and of stable polarity, whereas, as the driving is increased and inertial effects play a more important part, the dynamos become multipolar and reversing^{29–32}. In these previous studies, the dynamos have viscous and Ohmic diffusivities that are broadly similar ($Pr_m = \mathcal{O}(1)$). S6, with its small Pr_m and Ek , has $Ro_l = 0.06$ (Extended Data Table 2) and so contradicts the accepted regime division by lying in the purportedly stable dipolar regime. Similarly, our dynamo S6 should lie close to the non-reversing regime ($Ro_l < 0.05$) according to a study of magnetic reversal frequency scaling⁸. Conversely, it exhibits a reversal frequency that is predicted to occur only at $Ro_l = 0.4$ according to dynamos for which $Ek > 3 \times 10^{-4}$ and $3 \leq Pr_m \leq 20$.

The polarity reversals exhibited by S6 are nearly periodic, but if the Rayleigh number is increased, for example, as in S6.07, the time dependence becomes more variable, and Earth-like events such as dipole excursions become possible (Extended Data Fig. 2). Moreover, at a radius corresponding to Earth’s surface, the field in S6 is dipole-dominated between reversals (Supplementary Video 7). Our results therefore indicate that dynamo waves could indeed play a part in long-term geomagnetic variability^{2,33}. Unfortunately, it is currently a huge computational undertaking to produce the long time series needed to fully characterize the statistics of reversals in the low- Ek regime: S6 required 4×10^6 CPU hours (approximately 457 CPU years) to integrate for 6.5 magnetic diffusion times. More complete studies of the dynamics of polarity reversals in dynamos at low viscosity may therefore have to wait for new numerical tools that are capable of more efficiently exploring this challenging regime.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 1 February; accepted 2 September 2016.

Published online 7 November 2016.

1. Moffatt, H. K. *Field Generation in Electrically Conducting Fluids* (Cambridge Univ. Press, 1978).
2. Parker, E. N. *Cosmical Magnetic fields: Their Origin and Their Activity* (Oxford Univ. Press, 1979).
3. Jacobs, J. A. *Reversals of the Earth’s Magnetic Field* (Cambridge Univ. Press, 1994).
4. Merrill, R. T. & McFadden, P. L. Geomagnetic polarity transitions. *Rev. Geophys.* **37**, 201–226 (1999).

5. P  tr  lis, F. & Fauve, S. Mechanisms for magnetic field reversals. *Philos. Trans. R. Soc. Lond. A* **368**, 1595–1605 (2010).
6. Olson, P., Driscoll, P. & Amit, H. Dipole collapse and reversal precursors in a numerical dynamo. *Phys. Earth Planet. Inter.* **173**, 121–140 (2009).
7. Christensen, U. & Aubert, J. Scaling properties of convection-driven dynamos in rotating spherical shells and application to planetary magnetic fields. *Geophys. J. Int.* **166**, 97–114 (2006).
8. Olson, P. & Amit, H. Magnetic reversal frequency scaling in dynamos with thermochemical convection. *Phys. Earth Planet. Inter.* **229**, 122–133 (2014).
9. Parker, E. N. Hydromagnetic dynamo models. *Astrophys. J.* **122**, 293–314 (1955).
10. Busse, F. H. & Simitev, R. D. Parameter dependences of convection-driven dynamos in rotating spherical fluid shells. *Geophys. Astrophys. Fluid Dyn.* **100**, 341–361 (2006).
11. Christensen, U. R. & Wicht, J. In *Treatise on Geophysics* 2nd edn (ed. Schubert, G.) Vol. 8, 245–277 (Elsevier, 2015).
12. Roberts, P. H. & King, E. M. On the genesis of the Earth's magnetism. *Rep. Prog. Phys.* **76**, 096801 (2013).
13. Kida, S., Araki, K. & Kitauchi, H. Periodic reversals of magnetic field generated by thermal convection in a rotating spherical shell. *J. Phys. Soc. Jpn* **66**, 2194–2201 (1997).
14. Sarson, G. R. Reversal models from dynamo calculations. *Philos. Trans. R. Soc. Lond. A* **358**, 921–942 (2000).
15. Amit, H., Leonhardt, R. & Wicht, J. Polarity reversals from paleomagnetic observations and numerical dynamo simulations. *Space Sci. Rev.* **155**, 293–335 (2010).
16. Wicht, J. & Olson, P. A detailed study of the polarity reversal mechanism in a numerical dynamo model. *Geochem. Geophys. Geosyst.* **5**, Q03H10 (2004).
17. Glatzmaier, G. A. Geodynamo simulations—how realistic are they? *Annu. Rev. Earth Planet. Sci.* **30**, 237–257 (2002).
18. Busse, F. & Simitev, R. Toroidal flux oscillation as possible cause of geomagnetic excursions and reversals. *Phys. Earth Planet. Inter.* **168**, 237–243 (2008).
19. Roberts, P. H. Kinematic dynamo models. *Philos. Trans. R. Soc. Lond. A* **272**, 663–698 (1972).
20. Gubbins, D. & Gibbons, S. Three-dimensional dynamo waves in a sphere. *Geophys. Astrophys. Fluid Dyn.* **96**, 481–498 (2002).
21. Sakuraba, A. & Roberts, P. H. Generation of a strong magnetic field using uniform heat flux at the surface of the core. *Nat. Geosci.* **2**, 802–805 (2009).
22. Sakuraba, A. & Roberts, P. H. In *The Earth's Magnetic Interior* (eds Petrovsk  y, E. et al.) 117–129 (Springer, 2011).
23. Coe, R. S. & Glatzmaier, G. A. Symmetry and stability of the geomagnetic field. *Geophys. Res. Lett.* **33**, L21311 (2006).
24. P  tr  lis, F., Fauve, S., Dormy, E. & Valet, J.-P. Simple mechanism for reversals of Earth's magnetic field. *Phys. Rev. Lett.* **102**, 144503 (2009).
25. Gastine, T., Duarte, L. & Wicht, J. Dipolar versus multipolar dynamos: the influence of the background density stratification. *Astron. Astrophys.* **546**, A19 (2012).
26. Jones, C. A dynamo model of Jupiter's magnetic field. *Icarus* **241**, 148–159 (2014).
27. Dietrich, W., Schmitt, D. & Wicht, J. Hemispherical Parker waves driven by thermal shear in planetary dynamos. *Europhys. Lett.* **104**, 49001 (2013).
28. Dub  , C. & Charbonneau, P. Stellar dynamos and cycles from numerical simulations of convection. *Astrophys. J.* **775**, 69 (2013).
29. Kutzner, C. & Christensen, U. From stable dipolar towards reversing numerical dynamos. *Phys. Earth Planet. Inter.* **131**, 29–45 (2002).
30. Olson, P. & Christensen, U. R. Dipole moment scaling for convection-driven planetary dynamos. *Earth Planet. Sci. Lett.* **250**, 561–571 (2006).
31. Schirmer, M., Petitdemange, L. & Dormy, E. Dipole collapse and dynamo waves in global direct numerical simulations. *Astrophys. J.* **752**, 121 (2012).
32. Oruba, L. & Dormy, E. Transition between viscous dipolar and inertial multipolar dynamos. *Geophys. Res. Lett.* **41**, 7115–7120 (2014).
33. Gubbins, D. Mechanism for geomagnetic polarity reversals. *Nature* **326**, 167–169 (1987).

Supplementary Information is available in the online version of the paper.

Acknowledgements We acknowledge the computational resources provided by the Centro Svizzero di Calcolo Scientifico (CSCS) under the project s577. We are grateful to J. M. Favre from CSCS for assistance with high-performance visualization. We thank A. Willis for developing the original dynamo code that was used for the calculations reported here, and P. Marti for subsequent optimizations for the CSCS Cray. This work was partially supported by ERC grant no. 247303 (MFECE) to A.J. and by the Danish Council for Independent Research (DFR) grant no. 4002-00366 to C.C.F.

Author Contributions A.S. carried out the numerical geodynamo simulations and analysed the runs, C.C.F. drafted the manuscript and participated in the analysis of the results. All authors contributed equally to the design of the study, discussed the results, and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.S. (andrei.sheiko@erdw.ethz.ch).

Reviewer Information *Nature* thanks P. Olson and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

Governing equations and non-dimensionalization. We adopt the Boussinesq approximation for convection-driven, rotating magnetohydrodynamics, which results in the following non-dimensional equations:

$$\begin{aligned} \left(\text{Ro} \frac{\partial}{\partial t} - \text{Ek} \nabla^2 \right) \mathbf{u} &= \mathbf{N}_u - \nabla \hat{P} \\ \left(\frac{\partial}{\partial t} - \nabla^2 \right) \mathbf{B} &= \nabla \times (\mathbf{u} \times \mathbf{B}) \\ \left(\frac{\partial}{\partial t} - q \nabla^2 \right) T &= \varepsilon - \mathbf{u} \cdot \nabla T \end{aligned}$$

where

$$\mathbf{N}_u = \text{Ro} \mathbf{u} \times (\nabla \times \mathbf{u}) + (\nabla \times \mathbf{B}) \times \mathbf{B} + q \text{Ra} T \mathbf{r} - \hat{\mathbf{z}} \times \mathbf{u}$$

The variables \mathbf{u} , \mathbf{B} and T are the velocity, magnetic field and temperature, respectively. $\nabla \hat{P}$ is the modified pressure that contains information about conservative forces. The axis of rotation of the system is z and $\hat{\mathbf{z}}$ is a unit vector in that direction. Time is denoted as t . A uniform heat source ε is included. Solenoidal conditions $\nabla \cdot \mathbf{B} = 0$ and $\nabla \cdot \mathbf{u} = 0$ are integrated into the solution technique through use of a poloidal-toroidal decomposition of the vector field. The non-dimensional parameters are

$$\begin{aligned} \text{Ro} &= \eta / (2\Omega d^2) \\ \text{Ek} &= \nu / (2\Omega d^2) \\ \text{Ra} &= g\alpha \Delta T d / (2\Omega \kappa) \\ q &= \kappa / \eta \end{aligned} \quad (2)$$

where Ro is the magnetic Rossby number, Ek is the Ekman number, Ra is the modified Rayleigh number and q is the Roberts number.

The units of length, time, magnetic field and temperature for the non-dimensional governing equations are chosen as

$$r \rightarrow dr, \quad t \rightarrow (d^2/\eta)t, \quad \mathbf{B} \rightarrow (2\Omega \rho_0 \mu_0 \eta)^{1/2} \mathbf{B}, \quad T \rightarrow \Delta T, \quad d = r_o - r_i$$

The following symbols denote the parameters of the system: $\Omega = \Omega \hat{\mathbf{z}}$ is the rotation rate, μ_0 is the permeability of free space, ρ_0 is the density, ν , κ and η are the kinematic viscosity, thermal diffusivity and magnetic diffusivity, respectively, α is the thermal expansivity and d is the thickness of the spherical shell of convecting fluid. The unit of temperature ΔT is chosen to be $\beta d / \beta^*$, where β and $\beta^* = -2(1 - r_i/r_o)$ are the dimensional and non-dimensional temperature gradients at the core–mantle boundary (CMB), respectively, and $r_{i,o}$ are the inner and outer radii of the shell, respectively. Gravity is assumed to vary linearly with radius and has value g on the outer boundary. The spherical coordinates are denoted (r, θ, φ) .

Boundary conditions and internal heating. The modelled fluid is enclosed in a rotating spherical shell between radii r_i and r_o with $c = r_i/r_o = 0.35$. Both boundaries are no-slip and impermeable. The outer boundary is electrically insulating, the inner core has the same electrical conductivity as the outer core. In case S6_InsIC the inner core is insulating. The inner-core temperature is kept constant at $T = \frac{c^2 + c + 2}{2(1 - c)c}$; the gradient of the temperature on the outer boundary is $\beta^* = -2/(1 - c)$. A uniform heat source with $\varepsilon = 3q$ is adopted throughout the outer core.

In run S6_{ε0}, the uniform heat source ε is set to zero.

Diagnostics. The dipole latitude is calculated as $\arcsin \left[g_1^0 / \sqrt{(g_1^0)^2 + (g_1^1)^2 + (h_1^1)^2} \right]$,

where $\{g_1^0, g_1^1, h_1^1\}$ are the first three Gauss coefficients of the magnetic field at the outer boundary. The kinetic and magnetic energies are defined as $E_{\text{kin}} = \frac{1}{2} \int u^2 dV$ and $E_{\text{mag}} = \frac{1}{2\text{Ro}} \int B^2 dV$, where the volume integral is over the entire outer core. The local Rossby number⁸ is defined as $\text{Ro}_l = 2 \frac{l_u}{\pi} \text{RoRe}_m$; the magnetic Reynolds number Re_m is Ud/η ; $l_u = \frac{\sum u_l^2}{2E_{\text{kin}}}$ is the mean harmonic degree; $u_l^2/2$ is the kinetic energy per harmonic degree l ; $U = \sqrt{2E_{\text{kin}}/V}$; and V is the volume of the spherical shell.

Numerical set-up. We solve the governing equations using a parameterization in spherical harmonics up to degree and order 255 for the angular component and 528 finite-difference points in radius. A second-order predictor–corrector scheme is used for the time integration³⁴. The time-step is adaptive and varies throughout the run. In the case S6, the time-step started from 10^{-7} magnetic diffusion times at the initial stage of the run with a very turbulent flow, and later increased up to 2.5×10^{-6} magnetic diffusion times during the reversing interval. Parallelization is carried out in radius. In the linear parts of the code, data are split over the spherical harmonics. 528 cores were used simultaneously for one simulation.

The bulk of the simulations and visualizations were performed on the super-computer Piz Daint (Cray XC 30) at the Swiss National Supercomputing Centre. The code was originally developed by Willis³⁵ and then subsequently optimized for the Cray XC 30 and successfully benchmarked against other dynamo codes³⁶.

Importance of the heating mode. We carried out an additional simulation S6_{ε0} to test how the particular choice of the internal homogeneous heating influences the reversing behaviour. In the two runs S6 and S6_{ε0}, the internal heating ε is set to $3q$ and 0, respectively. Other parameters are the same. The run S6_{ε0} was started from S6. As a consequence of changing the heating mode, the temperature drop between the inner core boundary (ICB) and the CMB is modified (see Extended Data Fig. 3). An important transient effect is the secular cooling that can be defined as the decrease in the thermal energy in the shell per unit time: $Q_{\text{sec}} = - \int_V \frac{\partial T}{\partial t} dV$

(to evaluate Q_{sec} , the mean radial temperature gradient on the ICB is measured; see Equation (5) and Extended Data Fig. 4). We plot the ratio Q_{sec}/Q_i^0 of the secular cooling to the heat coming from the ICB in the steady ($Q_{\text{sec}} = 0$) state in Extended Data Fig. 5. In the modified run S6_{ε0}, Q_{sec} constitutes less than 5% of the heat flux Q_i^0 from the ICB during the final magnetic diffusion time. So, we conclude that the contribution of the secular cooling term to the heat flux equation is small at this stage.

The reversal period in S6_{ε0} increased by less than 5% in comparison with that in S6 (see dipole latitude time dependencies for both runs in Extended Data Fig. 6). This test suggests that the inclusion or not of the internal heating term is not crucial to the dynamo-wave reversals exhibited by our dynamos. Changes in the internal heating term do nevertheless slightly alter the value of the reversal period obtained.

Volume-integrated terms of the heat equation. The heat flux equation is

$$\left(\frac{\partial}{\partial t} - q \nabla^2 \right) T = \varepsilon - \mathbf{u} \cdot \nabla T$$

In a quasi-stationary state $\int_V \frac{\partial T}{\partial t} dV = 0$, if no-slip boundary conditions ($u_\varphi = u_\theta = 0$) are applied, then

$$\int_V \mathbf{u} \cdot \nabla T dV = \int_{\partial V} \mathbf{u} T \cdot d\mathbf{S} = 0$$

and the integrated heat flux equation is

$$-q \int_{r_i} \frac{\partial T}{\partial r} dS + \varepsilon V = -q \int_{r_o} \frac{\partial T}{\partial r} dS$$

This is equivalent to $Q_i + Q_{\text{int}} = Q_o$, whereby the sum of the heat supplied by internal heating Q_{int} and that from the inner core Q_i is equal to the heat leaving through the outer boundary Q_o .

The temperature gradient at the CMB was held fixed at $\beta^* = \frac{\partial T}{\partial r} \Big|_{r_o} = -2/(1 - c)$, with $c = r_i/r_o = 0.35$. In S6 the internal heating $Q_{\text{int}} = 3qV$. The heat leaving through the outer boundary $Q_o = -4q\pi r_o^2 \beta^*$. The temperature gradient on the ICB in run S6 is

$$\frac{\partial T}{\partial r} \Big|_{r_i} = -(Q_o - Q_{\text{int}})/(4q\pi r_i^2) = \frac{\beta^*}{c^2} + \left(\frac{1}{c^3} - 1 \right) r_i = -13.1 \quad (3)$$

In the case S6_{ε0} without internal heating, the temperature gradient at r_i is

$$\frac{\partial T}{\partial r} \Big|_{r_i} = -Q_o/(4q\pi r_i^2) = \frac{\beta^*}{c^2} = -25.1 \quad (4)$$

Without internal heating, in a steady case the heat from the ICB goes through the shell and leaves without losses or gains at the CMB. The power flowing through the CMB and the ICB is equal, although the heat flux is smaller at the CMB because it is inversely proportional to the surface area, yielding the coefficient $(r_o/r_i)^2 = 1/c^2$ in equation (4).

Let us consider the non-steady case:

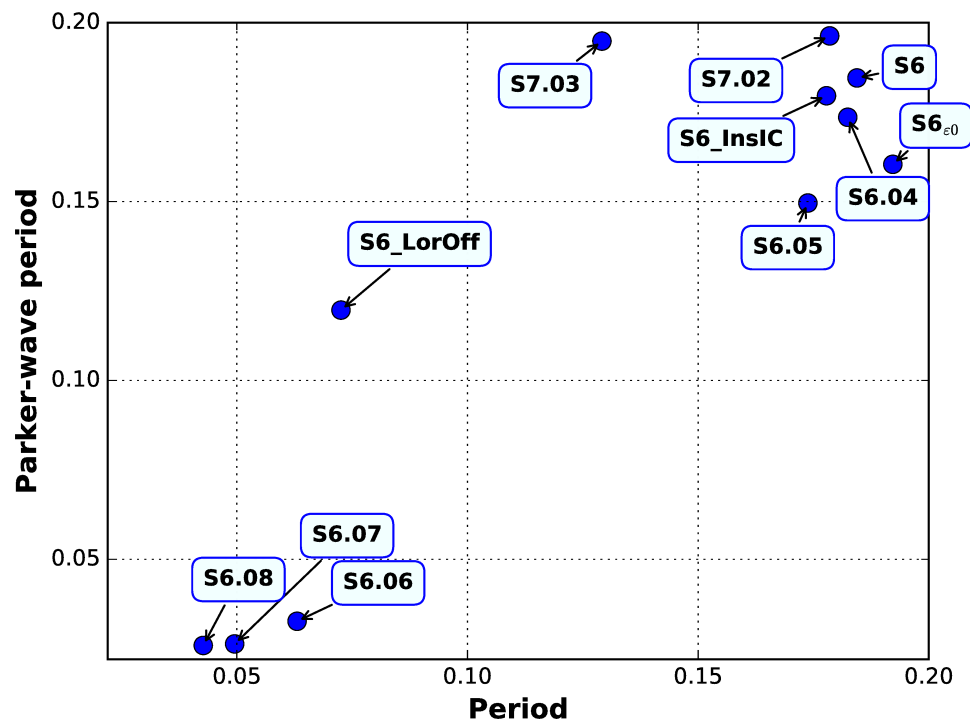
$$Q_i + Q_{\text{int}} = Q_o + \int_V \frac{\partial T}{\partial t} dV$$

where $Q_{\text{sec}} = - \int_V \frac{\partial T}{\partial t} dV$ is the secular cooling. If the heat flux from the inner core

in the steady case is Q_i^0 , the $Q_{\text{sec}} = Q_i^0 - Q_i$. So, the difference between the steady state and the actual value of the heat flux from the inner boundary is the secular cooling. The relative influence of the secular cooling can be calculated as

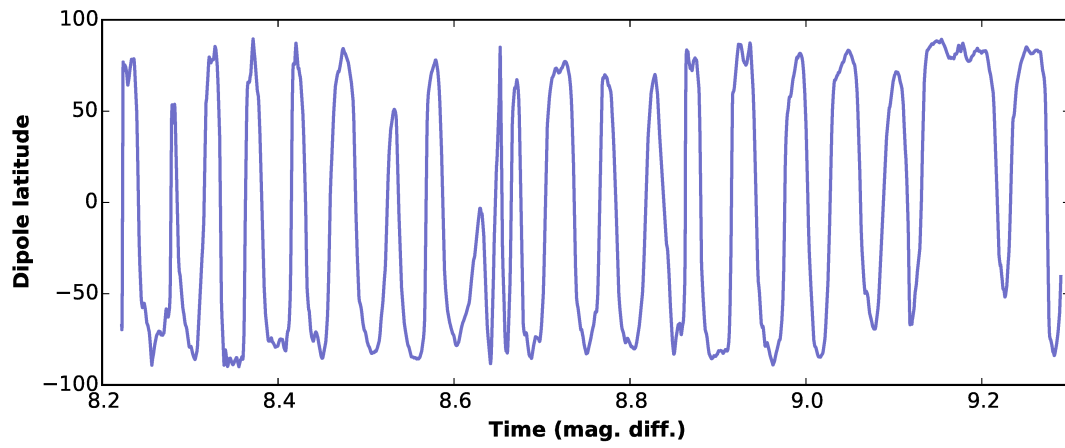
$$Q_{\text{sec}}/Q_i^0 = 1 - Q_i/Q_i^0 \quad (5)$$

34. Sheyko, A. *Numerical Investigations of Rotating MHD in a Spherical Shell*. PhD thesis, ETH Zürich, 24–26 (2014); <http://e-collection.library.ethz.ch/view/eth:9016>.
35. Willis, A. P., Sreenivasan, B. & Gubbins, D. Thermal core–mantle interaction: exploring regimes for locked dynamo action. *Phys. Earth Planet. Inter.* **165**, 83–92 (2007).
36. Jackson, A. *et al.* A spherical shell numerical dynamo benchmark with pseudo-vacuum magnetic boundary conditions. *Geophys. J. Int.* **196**, 712–723 (2014).
37. Rotvig, J. An investigation of reversing numerical dynamos driven by either differential or volumetric heating. *Phys. Earth Planet. Inter.* **176**, 69–82 (2009).
38. Driscoll, P. & Olson, P. Effects of buoyancy and rotation on the polarity reversal frequency of gravitationally driven numerical dynamos. *Geophys. J. Int.* **178**, 1337–1350 (2009).
39. Takahashi, F., Matsushima, M. & Honkura, Y. A numerical study on magnetic polarity transition in an MHD dynamo model. *Earth Planets Space* **59**, 665–673 (2007).
40. Olson, P. & Deguen, R. Eccentricity of the geomagnetic dipole caused by lopsided inner core growth. *Nat. Geosci.* **5**, 565–569 (2012).
41. Wicht, J., Stellmach, S. & Harder, H. in *Geomagnetic Field Variations* (eds Glaßmeier, K.-H. *et al.*) 107–158 (Springer, 2009).
42. Aubert, J., Aurnou, J. & Wicht, J. The magnetic structure of convection-driven numerical dynamos. *Geophys. J. Int.* **172**, 945–956 (2008).
43. Sreenivasan, B., Sahoo, S. & Dhama, G. The role of buoyancy in polarity reversals of the geodynamo. *Geophys. J. Int.* **199**, 1698–1708 (2014).

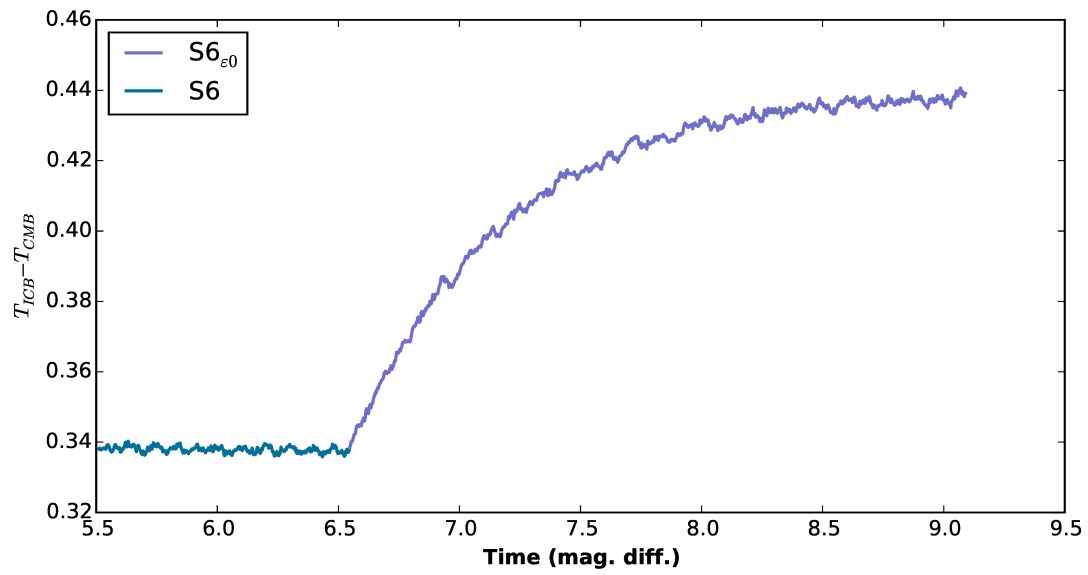


Extended Data Figure 1 | Systematics of reversal period versus dynamo-wave predictions. Comparison of the scaling of periods predicted by Parker's dynamo-wave theory (Equation (1)) and the periods obtained in the dynamo calculations. Periods in the numerical calculations are determined from spectral analysis of the time series of dipole latitudes. The predicted periods from Parker's dynamo-wave theory are calculated

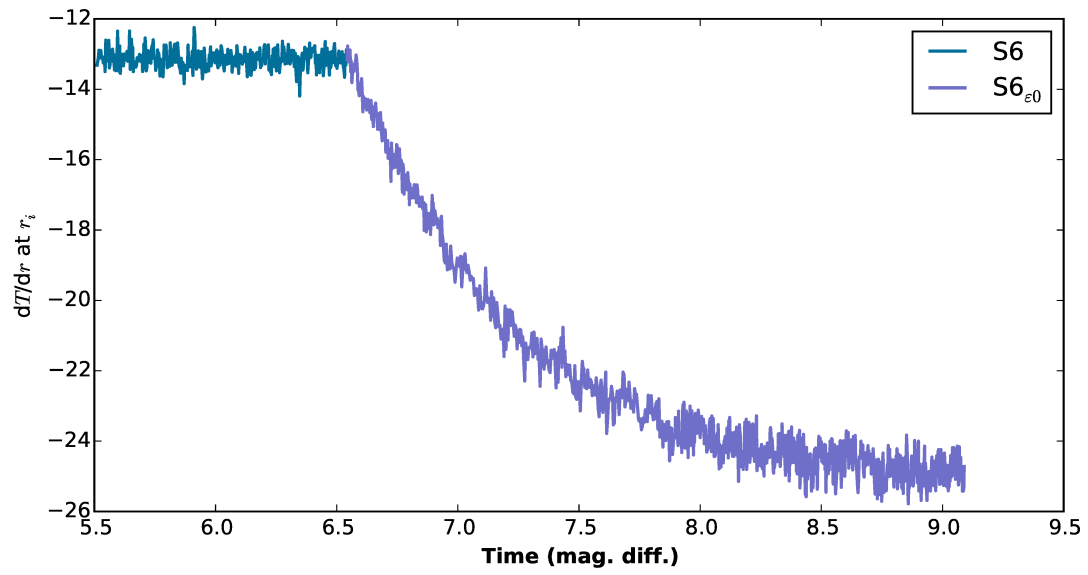
from volume-averaged helicity and zonal toroidal kinetic energy time-averaged over the last three reversals in the simulation or, in shorter simulations, during the second half of the reversing interval, with $\gamma = 200$ (Equation (1)). Numerical values of measured and estimated periods are provided in Extended Data Table 2.



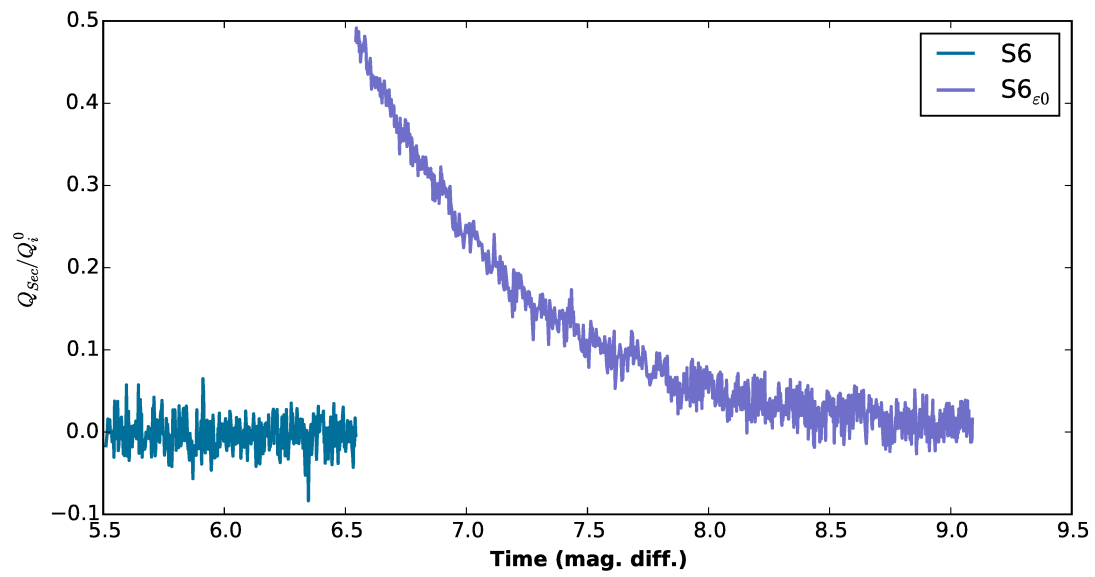
Extended Data Figure 2 | Time dependence of dipole tilt for dynamo S6.07. The time dependence of the dipole tilt, determined from the first three Gauss coefficients $\{g_1^0, g_1^1, h_1^1\}$ of the magnetic field at the outer boundary. Units are magnetic diffusion times. S6.07 was started from a previous run S6.06, and then the Rayleigh number Ra was increased by a factor of two.



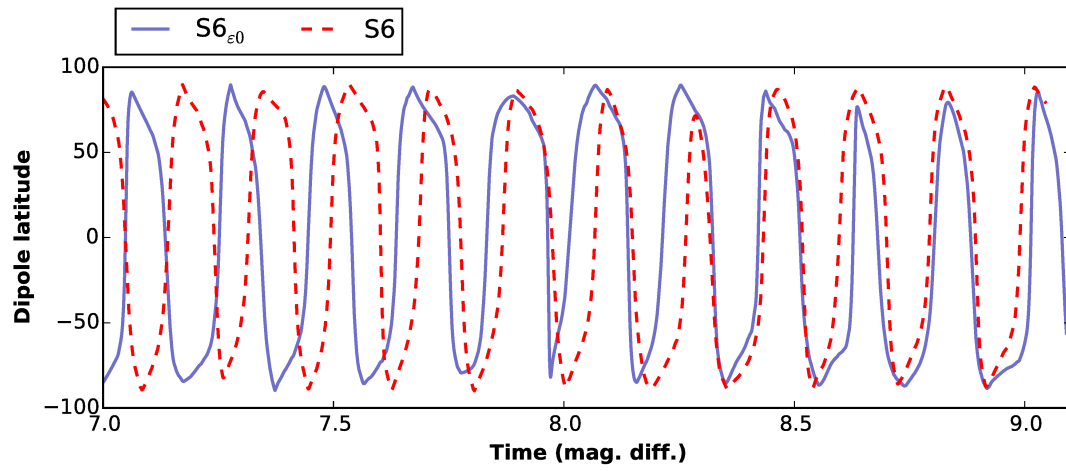
Extended Data Figure 3 | The temperature drop between the ICB and the CMB in runs S6 and S6_{ε0}. Run S6_{ε0} starts from S6; the temperature drop increases after the internal heating is switched off.



Extended Data Figure 4 | Radial temperature gradient on the ICB in runs S6 and S6_{ε0}. Steady-state values are -13.1 and -25.1 for runs S6 and S6_{ε0}, respectively (see Equations (3) and (4)).



Extended Data Figure 5 | Secular cooling normalized by the steady-state heat flux through the CMB (see Equation (5)), for runs S6 and S6_{ε0}.



Extended Data Figure 6 | Time dependence of dipole latitude for dynamos $S6$ and $S6_{\epsilon_0}$. The plot for $S6$ is shifted along the time axis to overlap $S6_{\epsilon_0}$.

Extended Data Table 1 | Selection of previous geodynamo model reversal studies included in Fig. 1

Ek	$Pr_m = Ek/Ro$	Author
$\mathcal{O}(10^{-15})$	$\mathcal{O}(10^{-5})$	Earth
$10^{-3} - 10^{-4}$	0.67–6	Kutzner and Christensen (2002) ²⁹
$10^{-2} - 3 \times 10^{-4}$	20	Rotvig (2009) ³⁷
5×10^{-3}	20	Driscoll and Olson (2009) ³⁸
6.5×10^{-3}	20	Olson et al. (2009) ⁶
10^{-5}	0.5	Takahashi et al. (2007) ³⁹
10^{-2}	20	Wicht and Olson (2004) ¹⁶
6×10^{-3}	20	Olson and Deguen (2012) ⁴⁰
$2 \times 10^{-2} - 3 \times 10^{-4}$	3–10	Wicht et al. (2009) ⁴¹
10^{-5}	0.5	Busse and Simitev (2008) ¹⁸
$2 \times 10^{-2} - 3 \times 10^{-4}$	3–10	Aubert et al. (2008) ⁴²
1.1×10^{-2}	10	Kida et al. (1997) ¹³
3×10^{-4}	5	Sreenivasan et al. (2014) ⁴³
1.1834×10^{-6}	0.04–0.05	Present study

We list two of the important control parameters, the magnetic Prandtl number Pr_m and the Ekman number Ek . Our aim was to make Pr_m substantially less than unity and to decrease the smallest previously reported value for Ek by at least an order of magnitude. Values from previous studies are from refs 6, 13, 16, 18, 29, 37–43.

Extended Data Table 2 | Runs exploring variation of reversals with control parameters and set-up

	Ek	Ro	Ra	q	Re_m	Ro_l	E_{kin}/E_{mag}	$\overline{E_{tor}}$	H	T_{osc}	T_{Parker}
S6	1.1834e-06	2.3669e-05	6591	0.050	89	0.064	9.22	5.08e+03	6.51e+05	0.184	0.185
S6.04	1.1834e-06	2.3669e-05	7250	0.050	94	0.071	8.93	5.65e+03	6.98e+05	0.182	0.174
S6.05	1.1834e-06	2.3669e-05	10000	0.050	105	0.084	6.53	6.98e+03	8.45e+05	0.174	0.150
S6.06	1.1834e-06	2.3669e-05	100000	0.050	333	0.226	3.22	9.19e+04	4.88e+06	0.063	0.033
S6.07	1.1834e-06	2.3669e-05	200000	0.050	388	0.258	2.91	1.40e+05	6.08e+06	0.050	0.026
S6.08	1.1834e-06	2.3669e-05	250000	0.050	396	0.286	3.06	1.40e+05	6.30e+06	0.043	0.026
S7.02	1.1834e-06	2.6299e-05	6591	0.045	84	0.066	12.56	5.73e+03	5.41e+05	0.178	0.196
S7.03	1.1834e-06	2.9586e-05	6591	0.040	83	0.065	31.23	8.15e+03	4.61e+05	0.129	0.195
S6_InsIC	1.1834e-06	2.3669e-05	6591	0.050	91	0.066	10.27	5.57e+03	6.56e+05	0.178	0.180
S6_LorOff	1.1834e-06	2.3669e-05	6591	0.050	119	0.061	—	2.54e+04	6.92e+05	0.073	0.120
S6 $_{\varepsilon 0}$	1.1834e-06	2.3669e-05	6591	0.050	100	0.076	6.12	6.00e+03	7.92e+05	0.192	0.160

T_{osc} is the period of reversals determined from spectral analysis of the model dipole tilt time series. T_{Parker} is the period of reversals determined by Equation (1) with $\gamma=200$. $\overline{E_{tor}}$ is the energy of the axially symmetric component of the toroidal flow. H is the modulus of the volume-integrated kinetic helicity of the non-axisymmetric flow in the northern hemisphere. Definitions of other parameters and diagnostics are provided in Methods. Presented diagnostics are averaged over the last three reversals or, in shorter simulations, during the second half of the reversing interval.

A cannabinoid link between mitochondria and memory

Etienne Hebert^{1,2,3*}, Tiffany Desprez^{1,2*}, Román Serrat^{1,2*}, Luigi Bellocchio^{1,2,4*}, Edgar Soria-Gomez^{1,2}, Arnau Busquets-Garcia^{1,2}, Antonio Christian Pagano Zottola^{1,2}, Anna Delamarre^{1,2}, Astrid Cannich^{1,2}, Peggy Vincent^{1,2}, Marjorie Varilh^{1,2}, Laurie M. Robin^{1,2}, Geoffrey Terral^{1,2}, M. Dolores García-Fernández^{5,6}, Michelangelo Colavita^{1,2,7}, Wilfrid Mazier^{1,2}, Filippo Drago⁷, Nagore Puente^{8,9}, Leire Reguero^{8,9}, Izaskun Elezgarai^{8,9}, Jean-William Dupuy¹⁰, Daniela Cota^{1,2}, Maria-Luz Lopez-Rodriguez¹¹, Gabriel Barreda-Gómez⁵, Federico Massa^{1,2}, Pedro Grandes^{8,9,12}, Giovanni Bénard^{1,2§} & Giovanni Marsicano^{1,2§}

Cellular activity in the brain depends on the high energetic support provided by mitochondria, the cell organelles which use energy sources to generate ATP^{1–4}. Acute cannabinoid intoxication induces amnesia in humans and animals^{5,6}, and the activation of type-1 cannabinoid receptors present at brain mitochondria membranes (mtCB₁) can directly alter mitochondrial energetic activity^{7–9}. Although the pathological impact of chronic mitochondrial dysfunctions in the brain is well established^{1,2}, the involvement of acute modulation of mitochondrial activity in high brain functions, including learning and memory, is unknown. Here, we show that acute cannabinoid-induced memory impairment in mice requires activation of hippocampal mtCB₁ receptors. Genetic exclusion of CB₁ receptors from hippocampal mitochondria prevents cannabinoid-induced reduction of mitochondrial mobility, synaptic transmission and memory formation. mtCB₁ receptors signal through intra-mitochondrial Gα_i protein activation and consequent inhibition of soluble-adenylyl cyclase (sAC). The resulting inhibition of protein kinase A (PKA)-dependent phosphorylation of specific subunits of the mitochondrial electron transport system eventually leads to decreased cellular respiration. Hippocampal inhibition of sAC activity or manipulation of intra-mitochondrial PKA signalling or phosphorylation of the Complex I subunit NDUF52 inhibit bioenergetic and amnesic effects of cannabinoids. Thus, the G protein-coupled mtCB₁ receptors regulate memory processes via modulation of mitochondrial energy metabolism. By directly linking mitochondrial activity to memory formation, these data reveal that bioenergetic processes are primary acute regulators of cognitive functions.

Cannabinoids can activate CB₁ receptors localized at the plasma or mitochondrial membranes^{7–9}. Exogenous cannabinoids induce amnesia^{5,6}, but the role of specific subcellular pools of CB₁ receptors in these effects is unknown.

In silico subcellular protein localization analysis^{10,11} revealed that the deletion of the 22 N-terminal amino acids of the CB₁ protein reduced its theoretical probability of mitochondrial targeting from 40–45% to 1–3% (Extended Data Table 1). The CB₁-receptor agonists WIN55,212-2 (WIN) and HU210 decreased respiration in CB₁-transfected primary mouse fibroblasts that were deficient in CB₁ (also known as *Cnr1*) or HEK 293 cells but not in cells transfected with a mutant cDNA of CB₁, which lacked the first 22 amino acids

(DN22-CB₁; Fig. 1a and Extended Data Fig. 1a, b). DN22-CB₁ receptors were functional, because cannabinoid stimulation of CB₁- and DN22-CB₁-transfected cells induced phosphorylation of extracellular-signal-regulated kinases (ERKs) and activation of G proteins to an equal extent (Fig. 1b, c and Extended Data Fig. 1c).

In neurons, distal energetic support requires mitochondrial mobility², which is reduced by cannabinoids in enteric neurons¹². HU210 significantly reduced the percentage of mobile mitochondria in axons of CB₁-, but not DN22-CB₁-transfected hippocampal neurons from CB₁^{−/−} mice (Fig. 1d, Extended Data Fig. 1d and Supplementary Videos 1, 2).

Semi-quantification of immunogold electron microscopy images⁸ revealed that the injection of an adeno-associated virus expressing CB₁ (AAV-CB₁) into the hippocampus of CB₁^{−/−} mice (denoted hereafter as CB₁^{−/−}(CB₁) mice) induced the same proportion of mtCB₁ receptors as in control mice injected with an AAV expressing GFP (CB₁^{+/+}(GFP) mice; Fig. 1e, f). CB₁^{−/−}(DN22-CB₁) mice expressed comparable amounts of total CB₁ protein as CB₁^{−/−}(CB₁) mice (Extended Data Fig. 2a, b), but the proportion of mtCB₁ receptors was similar to negative control CB₁^{−/−}(GFP) littermates (Fig. 1e, f). [³⁵S]GTPγ binding assays in hippocampal extracts of CB₁^{−/−}(GFP), CB₁^{−/−}(CB₁) and CB₁^{−/−}(DN22-CB₁) mice revealed that CB₁ and DN22-CB₁ receptors were functional *ex vivo* to an equal extent (Fig. 1g and Extended Data Fig. 2c–e). Thus, DN22-CB₁ is excluded from mitochondria without losing functionality, representing a useful tool to study mtCB₁ receptors in brain processes.

CB₁-receptor agonists reduce excitatory synaptic transmission in hippocampal slices from wild-type animals, but not from CB₁^{−/−} mice¹³. This effect was rescued in slices from CB₁^{−/−}(CB₁), but not from CB₁^{−/−}(DN22-CB₁) mice (Fig. 1h). Post-training cannabinoid administration induces amnesia in the novel object recognition (NOR) task¹⁴. Injection of WIN impaired the NOR performance of CB₁^{+/+}(GFP) mice, but not of CB₁^{−/−}(GFP) littermates, without altering total exploration (Fig. 1i and Extended Data Fig. 2f). This effect was fully rescued in CB₁^{−/−}(CB₁), but not in CB₁^{−/−}(DN22-CB₁) mice (Fig. 1i and Extended Data Fig. 2f).

We subsequently investigated intra-mitochondrial mtCB₁ receptor signalling. The cannabinoids Δ⁹-tetrahydrocannabinol (THC) and WIN decreased mitochondrial respiration in hippocampal cell cultures, reducing cellular and mitochondrial ATP content (Extended Data

¹INSERM U1215, NeuroCentre Magendie, Bordeaux 33077, France. ²Université de Bordeaux, NeuroCentre Magendie, Bordeaux 33077, France. ³Department of Biology, Université de Moncton, Moncton, New-Brunswick E1A 3E9, Canada. ⁴Department of Biochemistry and Molecular Biology I, Complutense University, Madrid 28040, Spain. ⁵Department of Research and Development, IMG Pharma Biotech S.L., Derio 48160, Spain. ⁶Department of Pharmacology, Faculty of Medicine and Dentistry, University of the Basque Country UPV/EHU, Leioa 48940, Spain. ⁷Department of Biomedical and Biotechnological Sciences, Section of Pharmacology, University of Catania, Catania 95124, Italy. ⁸Department of Neurosciences, Faculty of Medicine and Nursing, University of the Basque Country UPV/EHU, Leioa 48940, Spain. ⁹Achucarro Basque Center for Neuroscience, Bizkaia Science and Technology Park, Building 205, Zamudio 48170, Spain. ¹⁰Université de Bordeaux, Centre Génomique Fonctionnelle, Plateforme Protéome, Bordeaux 33077, France. ¹¹Department of Organic Chemistry, Complutense University, Madrid 28040, Spain. ¹²Division of Medical Sciences, University of Victoria, Victoria, British Columbia V8W 2Y2, Canada.

*These authors contributed equally to this work.

§These authors jointly supervised this work.

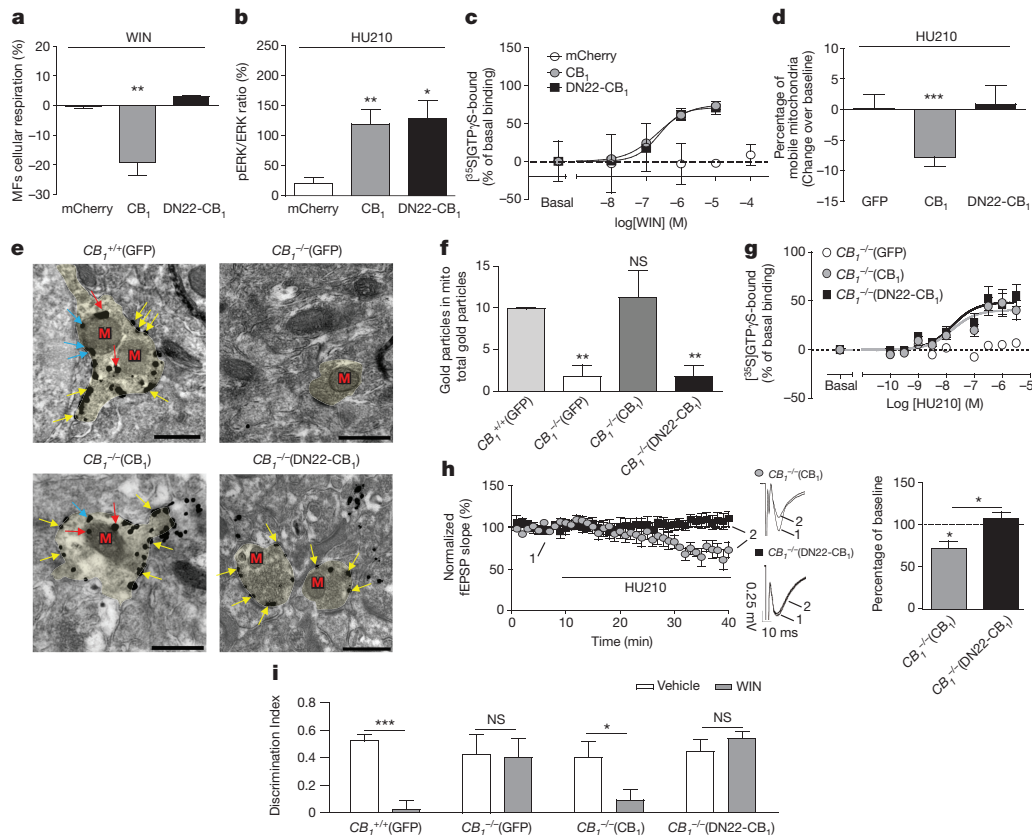


Figure 1 | Requirement of mtCB₁ receptors for cannabinoid effects on mitochondrial respiration and mobility, synaptic transmission and memory. **a**, The effect of WIN (2 μM) on cellular respiration of primary mouse fibroblasts (MFs) from *CB₁^{-/-}* pups (postnatal day (P)0–P1) transfected by electroporation with mCherry (mock)-, *CB₁*- or DN22-*CB₁*-expressing vectors (*n* = 3). **b**, The effect of HU210 (100 nM) on ERK phosphorylation in HEK293 cells transfected with mCherry- (*n* = 8), *CB₁*- (*n* = 13) or DN22-*CB₁*-expressing plasmids (*n* = 5). **c**, Dose–response curves of the effect of WIN on [³⁵S]GTPγS binding in membranes isolated from HEK293 cells transfected with mCherry- (*n* = 3–4), *CB₁*- or DN22-*CB₁*-expressing plasmids (*n* = 12–14). **d**, Effect of HU210 (1 μM) on the percentage of mobile mitochondria in axons from primary hippocampal neurons obtained from *CB₁^{-/-}* pups (P0–P1) and transfected with vectors expressing GFP (*n* = 17), *CB₁*-GFP (*n* = 32) or DN22-*CB₁*-GFP receptors (*n* = 14). **e**, Representative electromicrographs (3 independent experiments) of immunogold staining of CB₁ receptors in the CA1 hippocampal region of *CB₁^{+/+}* (GFP), *CB₁^{-/-}* (GFP), *CB₁^{-/-}* (*CB₁*), or *CB₁^{-/-}* (DN22-*CB₁*) mice. Scale bars, 0.5 μm. M, mitochondria; yellow arrows, plasma-membrane CB₁ receptors; red arrows, mtCB₁ receptors; blue arrows, excluded immunogold particles that were located on mitochondrial membranes but at a distance ≤80 nm from other

membranes¹⁰. **f**, Semi-quantification of the percentage of CB₁-positive immunogold particles (mtCB₁) in mitochondria (mito), normalized to the total number of gold particles (*n* = 3 mice, at least 20 images per mouse). **g**, Dose–response curves of the effect of HU210 on [³⁵S]GTPγS binding in membranes isolated from hippocampi of different mutant mice (*CB₁^{-/-}* (*CB₁*), *n* = 8; *CB₁^{-/-}* (DN22-*CB₁*), *n* = 15; *CB₁^{-/-}* (GFP), *n* = 22). **h**, The effect of HU210 (2.5 μM) on CA3–CA1 field excitatory post-synaptic potentials (fEPSP) slopes in hippocampal slices derived from *CB₁^{-/-}* (*CB₁*) (*n* = 5) and *CB₁^{-/-}* (DN22-*CB₁*) mice (*n* = 7). Left, plots of normalized fEPSP slopes with representative fEPSP traces before (1) and after HU210 treatment (2); right, histogram summarizing the average change in percentage of fEPSP slopes over baseline (dotted line). **i**, Cannabinoid-induced impairment of novel object recognition (NOR) memory requires mtCB₁-receptor activation in the hippocampus. Discrimination index values in the object-recognition test of *CB₁^{+/+}* (GFP) (*n* = 16), *CB₁^{-/-}* (GFP) (*n* = 7), *CB₁^{-/-}* (*CB₁*) (*n* = 13) and *CB₁^{-/-}* (DN22-*CB₁*) (*n* = 15) mice treated with vehicle (white bars) or the CB₁ agonist WIN (5 mg kg⁻¹, intraperitoneal (i.p.), grey bars). Data, mean ± s.e.m.; **P* < 0.05; ***P* < 0.01; NS, not significant. For statistics, see Supplementary Tables 1–3.

Fig. 3a, b). Mitochondrial G_α proteins^{15,16} co-immunoprecipitated with CB₁, and THC induced their release (Fig. 2a, b and Extended Data Fig. 3c). The G_αi inhibitor pertussis toxin (PTX) blocked the cannabinoid-induced decrease in mitochondrial cAMP, in protein kinase A (PKA) and complex I activity (Extended Data Fig. 3d–h) and in respiration (Fig. 2c, d).

sAC produces mitochondrial cAMP, thereby regulating respiration through PKA-mediated phosphorylation of oxidative phosphorylation (OXPHOS)-associated proteins¹⁷. As it was suggested that sAC acts independently of G proteins in the gonads¹⁸, we tested the prevalence of G_α–sAC interactions in brain mitochondria. In these organelles, sAC was detected along with other elements of the cAMP–PKA signalling cascade (Extended Data Fig. 3i) and it co-immunoprecipitated with G_α (Fig. 2e and Extended Data Fig. 4a). THC increased this interaction (Fig. 2e and Extended Data Fig. 4a), suggesting that sAC might mediate the mitochondrial effects of mtCB₁ receptors.

The activator of the non-mitochondrial membrane-bound AC, forskolin, had no effect on mitochondrial cAMP levels or on the decrease induced by the activation of mtCB₁ receptors (Extended Data Fig. 4b, c). Conversely, the sAC-activator bicarbonate (HCO₃⁻)¹⁸ fully reversed the effect of mtCB₁-receptor activation (Extended Data Fig. 4d, e) and abolished the cannabinoid-induced reduction in respiration (Fig. 2f, g). Overexpression of native or mitochondrial-enriched sAC protein¹⁷ (mtsAC; Fig. 2h and Extended Data Fig. 4g) in 3T3 cells expressing mtCB₁ and its mitochondrial signalling machinery (Extended Data Fig. 4f) abolished the effect of THC on cellular respiration (Fig. 2i). This effect was also present in hippocampal cultures (Extended Data Fig. 4h, i).

Next, we investigated whether sAC blockade prevents the effects of mtCB₁ receptors. The sAC inhibitor KH7 (refs 19, 20) slightly decreased respiration, but blocked the mtCB₁-dependent decrease in the activity of PKA and complex I (Extended Data Fig. 5a–c), and in respiration

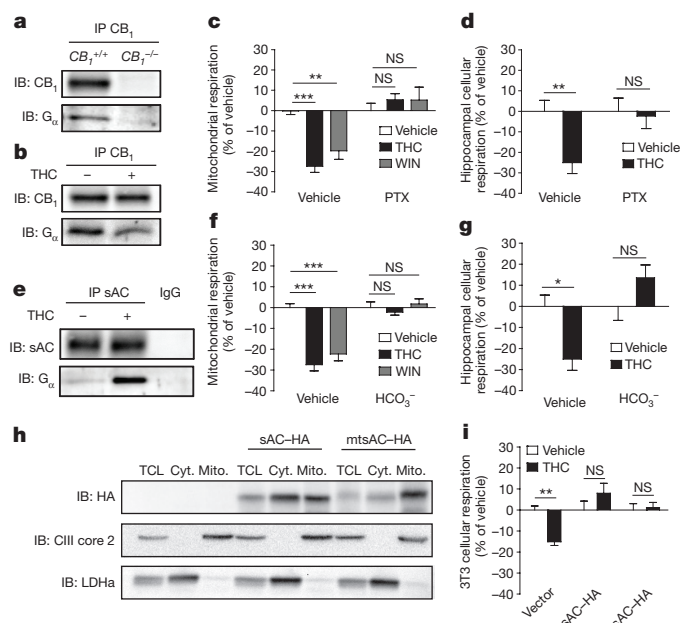


Figure 2 | Activation of mtCB₁ receptors affects mitochondrial activity via G_{αi/o} proteins and sAC. **a**, Representative immunoblotting (3 independent experiments) of co-immunoprecipitation experiments of G_α proteins with mtCB₁ receptors in purified brain mitochondria. **b**, Representative immunoblotting (3 independent experiments) showing the effect of Δ⁹-tetrahydrocannabinol (THC, 800 nM) on mtCB₁ and G_α proteins co-immunoprecipitation. **c**, The effect of pertussis toxin (PTX, 1 μg ml⁻¹) on the decrease in respiration in purified brain mitochondria induced by THC (800 nM) and WIN (100 nM) (*n* = 4–8). **d**, The effect of PTX on the decrease in cellular respiration in hippocampal cultures induced by THC (1 μM) (*n* = 8–9). **e**, Representative immunoblotting (3 independent experiments) of co-immunoprecipitation experiments of G_α proteins with sAC from purified brain mitochondria, after control or THC (800 nM) treatment. **f**, The effect of bicarbonate (HCO₃⁻, 5 mM) on the decrease in respiration in purified brain mitochondria induced by THC (800 nM) and WIN (100 nM) (*n* = 3–9). **g**, The effect of HCO₃⁻ on the decrease in cellular respiration in primary hippocampal cultures induced by THC (1 μM) (*n* = 8–11). **h**, Representative immunoblot (3 independent experiments) of haemagglutinin (HA), complex III core 2 (CIII core 2) and lactate dehydrogenase (LDHa) showing the presence of sAC-HA in total cellular (TCL), cytosolic (Cyt.) and mitochondrial (Mito) fractions, whereas mtsAC-HA is specifically targeted to mitochondrial fractions of transfected 3T3 cells. **i**, Overexpression of sAC-HA or mtsAC-HA in 3T3 cells blocks the decrease in respiration induced by THC (800 nM) (*n* = 6–8). IP, immunoprecipitation; IB, immunoblot. Data, mean ± s.e.m.; **P* < 0.05; ***P* < 0.01; ****P* < 0.001; NS, not significant. For statistics, see Supplementary Tables 1–3.

in brain mitochondria (Fig. 3a). KH7 or overexpression of mtsAC abolished the cannabinoid effect on neuronal mitochondrial mobility (Fig. 3b, Extended Data Fig. 5d, e and Supplementary Videos 3, 4). In hippocampal slices, KH7 blocked the cannabinoid-induced reduction of field excitatory post-synaptic potentials (fEPSPs) without altering synaptic transmission as such (Fig. 3c and Extended Data Fig. 5f, g). Eventually, the intrahippocampal infusion of KH7 abolished the amnesic effect of post-training WIN administration (Fig. 3d and Extended Data Fig. 5h).

Phosphorylation-dependent modulation of OXPHOS proteins^{17,21} might, therefore, mediate the effects of mtCB₁ receptors. The PKA activator 8-bromoadenosine 3',5'-cyclic monophosphate (8-Br-cAMP) reversed the cannabinoid-induced decrease in mitochondrial respiration, and the PKA-inhibitor H89 blocked the cannabinoid effects on the activity of mitochondrial PKA and complex I, and on respiration (Fig. 4a, b and Extended Data Fig. 6a–c). Silencing the A-kinase anchor protein AKAP121, which transports PKA to mitochondria^{22,23}, reduced intra-mitochondrial PKA levels and blocked the THC-induced effects

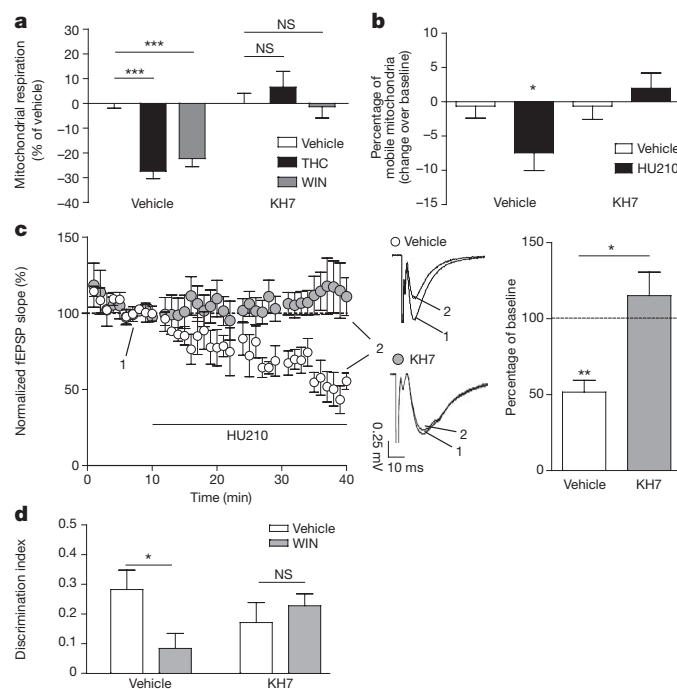


Figure 3 | Blockade of sAC activity prevents the effects of cannabinoids on mitochondrial respiration and mobility, synaptic transmission and memory. **a**, The effect of KH7 (5 μM) on the reduction in respiration in purified brain mitochondria induced by THC (800 nM) and WIN (100 nM) (*n* = 3–7). **b**, The effect of KH7 (5 μM) on the decrease in mitochondrial mobility in CB₁-expressing neurons induced by HU210 (1 μM) (*n* = 10–16). **c**, The effect of KH7 (10 μM) on the decrease in fEPSP at CA3–CA1 synapses in hippocampal slices from C57BL/6N mice induced by HU210 (2.5 μM) (*n* = 4–5). Left, plots of normalized fEPSP slopes with representative fEPSP traces before (1) and after (2) HU210 incubation. Vehicle (*n* = 4) or KH7 (*n* = 5) were pre-incubated 10 min before cannabinoid application. Right, histogram summarizing the average changes in percentage of fEPSP slope before (100% baseline, dotted line) and after HU210 treatment in presence or absence of KH7. **d**, Intrahippocampal KH7 blocks cannabinoid-induced amnesia in a NOR memory task. Discrimination index values in the NOR test showing the effect of vehicle or WIN (i.p., 5 mg kg⁻¹) in C57BL/6N mice that received intrahippocampal injection of vehicle or KH7 (2 μg in 0.5 μl). *n* = 7–14 per group. Data, mean ± s.e.m.; **P* < 0.05; ***P* < 0.01; ****P* < 0.001; NS, not significant. For statistics, see Supplementary Tables 1–3.

on complex I activity and respiration (Fig. 4c, d and Extended Data Fig. 6d–i). Systemic injection of THC reduced the PKA-dependent phosphorylation of brain mitochondrial proteins of wild-type, but not CB₁^{-/-} mice, and the intracerebroventricular injection of KH7 blocked this effect (Fig. 4e, f). These results indicate that mtCB₁ receptors modulate brain mitochondrial respiration through the PKA-dependent phosphorylation of mitochondrial proteins.

By using selective substrates, we observed a specific effect of THC on complex-I-dependent respiration in brain mitochondria (Extended Data Fig. 7a), suggesting that mtCB₁ receptor signalling specifically targets this complex. Analysis of the PKA-dependent phosphorylation of brain mitochondrial proteins revealed that the subunit NDUFS2 was consistently targeted by PKA in an mtCB₁-signalling-dependent manner (Fig. 4g–j and Extended Data Fig. 7b).

To study the functional role of the PKA–NDUFS2 signalling cascade, we first fused a constitutively active mutant form of PKA (PKA-CA)²⁴ to a mitochondrial leading sequence (MLS), inducing the mitochondrial localization of the mutant protein (MLS–PKA-CA; Extended Data Fig. 7c). Expression of MLS–PKA-CA inhibited the WIN-induced respiration decrease both in CB₁-co-transfected HEK293 cells (Extended Data Fig. 7d, e) and in hippocampal extracts from wild-type mice that received hippocampal injection of an AAV expressing the mutant

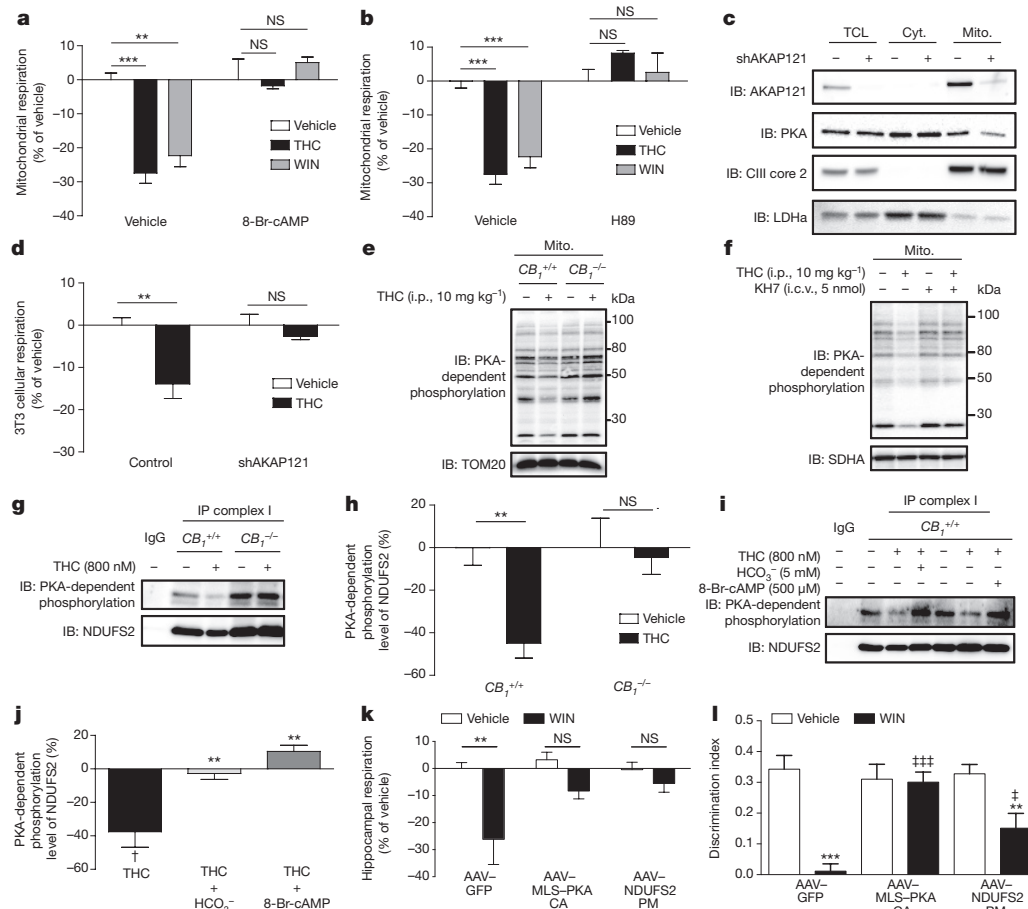


Figure 4 | Activation of mtCB₁ decreases PKA-dependent phosphorylation of mitochondrial proteins.

a, The effect of 8-Br-cAMP (500 μ M) on the reduction in respiration in purified brain mitochondria induced by THC (800 nM) and WIN (100 nM) ($n = 3-7$). **b**, The effects of H89 (800 nM) on the reduction in respiration in purified brain mitochondria induced by THC (800 nM) and WIN (100 nM) ($n = 3-7$). **c**, **d**, Silencing of AKAP121 (using small hairpin RNA, shAKAP121) lowers mitochondrial PKA and blocks the effect of mtCB₁-receptor activation. **c**, Representative immunoblots (3 independent experiments) of AKAP121, PKA, CIII core 2 and LDHa in total cellular, cytoplasmic and mitochondrial lysates from 3T3 cells treated with shAKAP121 or control. **d**, The effect of AKAP121 silencing on the reduction in cellular respiration in 3T3 cells induced by THC (800 nM) ($n = 4$). **e**, **f**, Activation of mtCB₁ decreases PKA-dependent phosphorylation of brain mitochondrial proteins. **e**, Representative immunoblots (3 independent experiments) showing the effect of THC (i.p., 10 mg kg⁻¹) on PKA-dependent phosphorylation levels of brain mitochondrial proteins ($n = 3$). TOM20, translocase of outer membrane subunit 20. **f**, Representative immunoblots (3 independent experiments) showing the effect of centrally-injected KH7 (intracerebroventricular injection (i.c.v.), 5 nmol) on the decrease

protein (AAV-MLS-PKA-CA; Fig. 4k and Extended Data Fig. 7f). The hippocampal expression of MLS-PKA-CA did not alter mouse memory *in vivo*, but it blocked the amnesic effect of WIN (Fig. 4l). We then generated a mutant form of NDUFS2, carrying point mutations at consensus sites for PKA-dependent phosphorylation^{25,26}, thereby providing a phosphomimetic NDUFS2 protein (NDUFS2-PM). NDUFS2-PM expression in CB₁-co-transfected HEK293 cells did not change cellular oxygen consumption, but it blocked the effect of WIN on cellular respiration (Extended Data Fig. 7c–e). The viral expression of NDUFS2-PM in mouse hippocampi (Extended Data Fig. 7f) inhibited the effect of WIN on respiration in hippocampal homogenates and substantially reduced the amnesic effect of the cannabinoid agonist (Fig. 4k, l).

These results show a previously unknown link between acute brain mitochondrial activity and memory, in which direct bioenergetic

in PKA-dependent phosphorylation of brain mitochondrial proteins induced by systemic THC injection (i.p., 10 mg kg⁻¹). SDHa, succinate dehydrogenase subunit A. **g–j**, Activation of mtCB₁ receptors decreases PKA-dependent phosphorylation of the complex I subunit NDUFS2. **g**, **h**, Representative immunoblotting (**g**; 3 independent experiments) and quantification (**h**; $n = 5$) of PKA-dependent phosphorylation and total NDUFS2 of immunoprecipitated complex I proteins from purified brain mitochondria treated with vehicle or THC (800 nM). **i**, **j**, Representative immunoblotting (**i**; 3 independent experiments) and quantification (**j**; $n = 3-7$) showing the effects of HCO₃⁻ (5 mM) and 8-Br-cAMP (500 μ M) on the decrease in PKA-dependent phosphorylation of NDUFS2 in brain mitochondria induced by THC (800 nM). **k**, The effect of WIN (100 nM) on respiration of homogenized hippocampi from mice that received intrahippocampal injection of AAV-MLS-PKA-CA or AAV-NDUFS2-PM ($n = 6$). **l**, The effect of WIN (5 mg kg⁻¹) on NOR discrimination index of mice intrahippocampally injected with AAV-MLS-PKA-CA or AAV-NDUFS2-PM ($n = 8-11$). Data, mean \pm s.e.m.; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; NS, not significant; † $P < 0.05$ compared to vehicle; ‡ $P < 0.05$; ‡‡‡ $P < 0.001$ compared to AAV-GFP WIN. For statistics, see Supplementary Tables 1–3.

effects of cannabinoids play a central role. The activation of intramitochondrial G α_i proteins by brain mtCB₁ receptors triggers inhibition of sAC activity, a reduction in cAMP levels and a decrease in PKA-dependent phosphorylation of OXPHOS proteins, among which the complex I subunit NDUFS2 plays a notable functional role. This cascade of events leads to a decrease in brain mitochondrial activity, which is required for the acute effects of cannabinoids on mitochondrial mobility, synaptic depression and eventually amnesia.

Mitochondria provide several necessary substrates for cells, having a particularly limiting role in brain functions^{1,2}. Whereas long-term mitochondrial alterations participate in the pathogenesis of neurological and neuropsychiatric diseases^{1,2} and can alter memory processes^{27,28}, recent evidence shows that even very short interruptions of the ATP supply exert marked effects on synaptic activity³.

The mtCB₁-dependent decrease in OXPHOS functions and ATP production represents a likely mechanism connecting mitochondrial functions and cannabinoid-induced synaptic and amnesic effects. More indirectly, recent work showed that the intracellular mTORC1 pathway participates in the cannabinoid-induced amnesic effects¹⁴. The brain mTORC1 complex bidirectionally interacts with mitochondria²⁹, suggesting that a common mitochondria–mTORC1 pathway might be involved in cannabinoid amnesic effects. Recent data revealed a biphasic effect of CB₁ receptor agonists in the hypothalamus, with low doses increasing and high doses decreasing respiration, respectively⁹. Thus, brain-region-specific or cell-type-specific differences in the signalling of mtCB₁ receptors might exist.

This study demonstrates that at least one G-protein-coupled receptor present on mitochondrial membranes modulates high brain functions such as memory formation through the modulation of intra-mitochondrial G-protein signalling. Considering that G proteins play a central role in the brain, the present data will probably pave the way for a new field of research that deals with the acute effects of mitochondrial activity on brain functioning.

Cannabinoid drugs have several therapeutic potentials³⁰, unfortunately limited by important side effects, such as impairment of memory^{5,6}. The present data suggest that selective targeting of specific subcellular populations of CB₁ receptors in the brain might assist in development of safer therapeutics against several brain disorders.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 7 October 2015; accepted 27 September 2016.

Published online 9 November 2016.

- Mattson, M. P., Gleichmann, M. & Cheng, A. Mitochondria in neuroplasticity and neurological disorders. *Neuron* **60**, 748–766 (2008).
- Sheng, Z. H. & Cai, Q. Mitochondrial transport in neurons: impact on synaptic homeostasis and neurodegeneration. *Nat. Rev. Neurosci.* **13**, 77–93 (2012).
- Rangaraju, V., Calloway, N. & Ryan, T. A. Activity-driven local ATP synthesis is required for synaptic function. *Cell* **156**, 825–835 (2014).
- Attwell, D. & Laughlin, S. B. An energy budget for signaling in the grey matter of the brain. *J. Cereb. Blood Flow Metab.* **21**, 1133–1145 (2001).
- Broyd, S. J., van Hell, H. H., Beale, C., Yücel, M. & Solowij, N. Acute and chronic effects of cannabinoids on human cognition—a systematic review. *Biol. Psychiatry* **79**, 557–567 (2016).
- Marsicano, G. & Lafenêtre, P. Roles of the endocannabinoid system in learning and memory. *Curr. Top. Behav. Neurosci.* **1**, 201–230 (2009).
- Bénard, G. et al. Mitochondrial CB₁ receptors regulate neuronal energy metabolism. *Nat. Neurosci.* **15**, 558–564 (2012).
- Hebert-Chatelain, E. et al. Cannabinoid control of brain bioenergetics: Exploring the subcellular localization of the CB₁ receptor. *Mol. Metab.* **3**, 495–504 (2014).
- Koch, M. et al. Hypothalamic POMC neurons promote cannabinoid-induced feeding. *Nature* **519**, 45–50 (2015).
- Claros, M. G. MitoProt, a Macintosh application for studying mitochondrial proteins. *Comput. Appl. Biosci.* **11**, 441–447 (1995).
- Nakai, K. & Horton, P. PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem. Sci.* **24**, 34–36 (1999).
- Boesmans, W., Ameloot, K., van den Abbeel, V., Tack, J. & Vanden Berghe, P. Cannabinoid receptor 1 signalling dampens activity and mitochondrial transport in networks of enteric neurones. *Neurogastroenterol. and Motil.* **21**, 958–968 (2009).
- Kano, M., Ohno-Shosaku, T., Hashimoto, Y., Uchigashima, M. & Watanabe, M. Endocannabinoid-mediated control of synaptic transmission. *Physiol. Rev.* **89**, 309–380 (2009).
- Puighearnal, E. et al. Cannabinoid modulation of hippocampal long-term memory is mediated by mTOR signaling. *Nat. Neurosci.* **12**, 1152–1158 (2009).
- Benincá, C. et al. A new non-canonical pathway of G_{αq} protein regulating mitochondrial dynamics and bioenergetics. *Cell. Signal.* **26**, 1135–1146 (2014).
- Lyssand, J. S. & Bajjalieh, S. M. The heterotrimeric G protein subunit G_{αi} is present on mitochondria. *FEBS Lett.* **581**, 5765–5768 (2007).
- Acin-Perez, R. et al. Modulation of mitochondrial protein phosphorylation by soluble adenylyl cyclase ameliorates cytochrome oxidase defects. *EMBO Mol. Med.* **1**, 392–406 (2009).
- Chen, Y. et al. Soluble adenylyl cyclase as an evolutionarily conserved bicarbonate sensor. *Science* **289**, 625–628 (2000).
- Pierre, S., Eschenhagen, T., Geisslinger, G. & Scholich, K. Capturing adenylyl cyclases as potential drug targets. *Nat. Rev. Drug Discov.* **8**, 321–335 (2009).
- Hess, K. C. et al. The ‘soluble’ adenylyl cyclase in sperm mediates multiple signaling events required for fertilization. *Dev. Cell* **9**, 249–259 (2005).
- Acin-Perez, R. et al. Cyclic AMP produced inside mitochondria regulates oxidative phosphorylation. *Cell Metab.* **9**, 265–276 (2009).
- Livigni, A. et al. Mitochondrial AKAP121 links cAMP and src signaling to oxidative metabolism. *Mol. Biol. Cell* **17**, 263–271 (2006).
- Chen, Q., Lin, R. Y. & Rubin, C. S. Organelle-specific targeting of protein kinase AII (PKAII). Molecular and *in situ* characterization of murine A kinase anchor proteins that recruit regulatory subunits of PKAII to the cytoplasmic surface of mitochondria. *J. Biol. Chem.* **272**, 15247–15257 (1997).
- Niswender, C. M. et al. Cre recombinase-dependent expression of a constitutively active mutant allele of the catalytic subunit of protein kinase A. *Genesis* **43**, 109–119 (2005).
- Amanchy, R. et al. A curated compendium of phosphorylation motifs. *Nat. Biotechnol.* **25**, 285–286 (2007).
- Neuberger, G., Schneider, G. & Eisenhaber, F. pKaPS: prediction of protein kinase A phosphorylation sites with the simplified kinase-substrate binding model. *Biol. Direct* **2**, 1–23 (2007).
- Lin, A. L. et al. Decreased *in vitro* mitochondrial function is associated with enhanced brain metabolism, blood flow, and memory in Surf1-deficient mice. *J. Cereb. Blood Flow Metab.* **33**, 1605–1611 (2013).
- Roubertoux, P. L. et al. Mitochondrial DNA modifies cognition in interaction with the nuclear genome and age in mice. *Nat. Genet.* **35**, 65–69 (2003).
- Bockaert, J. & Marin, P. mTOR in brain physiology and pathologies. *Physiol. Rev.* **95**, 1157–1187 (2015).
- Pacher, P., Bátkai, S. & Kunos, G. The endocannabinoid system as an emerging target of pharmacotherapy. *Pharmacol. Rev.* **58**, 389–462 (2006).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank D. Gonzales, N. Aubailly and all the personnel of the Animal Facility of the NeuroCentre Magendie, M. Biguerie of the technical service of the NeuroCentre Magendie, the personnel from the Bordeaux Imaging Center and V. Morales for continuous help. We thank G. Manfredi (Cornell University) for the sAC-HA, A. Feliciello (University of Napoli) for anti-AKAP121 antiserum and for AKAP121 shRNA, and M. Montcouquiol, N. Pigué (INSERM U1215, Bordeaux) and R. Rosignol (University of Bordeaux) for help with experiments. We thank F. Francia, P. V. Piazza, A. Bacci, G. Ferreira, F. Chaouloff and M. Guzman for critical reading and the members of Marsicano's laboratory for discussions. This work was supported by INSERM (G.M., D.C.), EU-Fp7 (PAINCAGE, HEALTH-603191, G.M. and FP7-PEOPLE-2013-IEF-623638, A.B.-G.), European Research Council (Endofood, ERC-2010-StG-260515 and CannaPreg, ERC-2014-PoC-640923, G.M.), Fondation pour la Recherche Médicale (DRM20101220445, G.M., SPF20121226369, R.S. and ARF20140129235, L.B.), Fondation pour la Recherche en Psychiatrie et en Santé Mentale (FRPSM, G.M.), Human Frontiers Science Program (RGP0036/2014, G.M.), Région Aquitaine (G.M.), AFM Telethon Trampoline Grant (16474, G.B.), Agence Nationale de la Recherche (ANR Blanc NeuroNutriSens ANR-13-BSV4-0006, G.M., D.C., BRAIN ANR-10-LABX-0043, G.M., D.C., F.M. and ANR-10-IDEX-03-02, A.B.-G.), Dulbecco Telethon Institute post-doc fellowship (E.H.-C.), NSERC (RGPIN-2015-05880, E.H.-C.), Fyssen Foundation (E.S.-G.), EMBO post-doc fellowship (L.B.), CONACYT (E.S.-G.), Zabalduz pre-doc fellowship (M.D.G.-F.), the Basque Government (IT764-13, P.G.), MINECO/FEDER (SAF2015-65034-R, P.G.), University of the Basque Country (UPV/EHU UFI11/41, P.G.), Red de Trastornos Adictivos—Instituto de Salud Carlos III (RD12/0028/0004, RD16/0017/0012, P.G.).

Author Contributions E.H.-C., T.D., R.S. and L.B. performed biochemical, molecular biology, behavioural and cellular experiments; E.S.-G., A.B.-G. and L.M.R. helped with behavioural experiments and analyses; E.S.-G., A.C.P.Z., A.C., A.D., P.V. and M.V. helped with biochemistry and molecular biology; G.T., M.C., F.D., W.M., D.C. and F.M. performed electrophysiological studies; M.D.G.-F. and G.B.-G. performed G protein signalling and binding experiments; N.P., L.R., I.E. and P.G. performed electron microscopy experiments; J.-W.D. provided proteomics experiments; M.-L.L.-R. provided reagents; G.B. and G.M. supervised the work; E.H.-C., T.D., G.B. and G.M. wrote the manuscript; all authors discussed results and edited the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to G.M. (giovanni.marsicano@inserm.fr).

Reviewer Information Nature thanks M. Mattson and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

Mice. Experiments were approved by the local ethical committee of the University of Bordeaux (approval number 501350-A) and the French Ministry of Agriculture and Forestry (authorization number 3306369). Mice were maintained under standard conditions (food and water ad libitum; 12 h–12 h light–dark cycle, light on at 7:00; experiments were performed between 9:00 and 17:00). Male C57BL/6N mice were purchased from Janvier (France). Wild-type ($CB_1^{+/+}$) and $CB_1^{-/-}$ female and male mice (2–4 months old) were obtained, bred and genotyped as described³¹. Only male mice were used for behavioural experiments. For most experiments $CB_1^{+/+}$ and $CB_1^{-/-}$ were littermates. For primary cell cultures, pups were obtained from homozygote pairs. No method of randomization to assign experimental groups was performed and the number of mice in each experimental group was similar. No statistical methods were used to predetermine sample size.

Drugs. THC was obtained from THC Pharm GmbH (Frankfurt, Germany). HU210 was synthesized as described³². WIN55-212-2, KH7, PTX, bicarbonate (HCO_3^-), forskolin, carbonyl cyanide-4-(trifluoromethoxy)phenylhydrazone (FCCP), oligomycin, antimycin, rotenone, picrotoxin, GTP γ S and other chemicals used in this study were purchased from Sigma-Aldrich (St-Louis, USA). [^3H] CP55,940 ($162.5 \text{ Ci mmol}^{-1}$) and [^3S]GTP γ S ($1,250 \text{ Ci mmol}^{-1}$) were purchased from Perkin Elmer NEN (Boston, USA).

For *in vivo* administration, WIN was dissolved in a mixture of saline (0.9% NaCl) with 2% DMSO and 2% cremophor; THC was dissolved in a mixture of 4% ethanol, 5% cremophor and saline; and KH7 was dissolved in 10% cremophor, 2.5% DMSO and saline. Vehicles contained the same amounts of solvents. All drugs were prepared freshly before the experiments.

For *in vitro* experiments, PTX, HCO_3^- and forskolin were dissolved in water. KH7, HU210 and WIN were dissolved in DMSO. THC, oligomycin, FCCP, antimycin and rotenone were dissolved in ethanol. Corresponding vehicle solutions were used in control experiments. DMSO and ethanol were no more than 0.001%.

Doses and concentrations of the different drugs were chosen on the basis of previous published data or preliminary experiments.

Plasmids. The N-terminal deletion of the first 22 amino acids (66 base pairs) in the mouse CB_1 -receptor coding sequence, to obtain the DN22- CB_1 mutant, as well as the generation of the mitochondrially targeted constitutively active form of PKA (MLS-PKA-CA) was achieved by polymerase chain reaction (PCR).

In brief, for DN22- CB_1 a forward primer hybridizing from the 67th base starting from ATG was coupled to a reverse primer hybridizing to the end of the coding sequence, including the TGA stop codon. In order to guarantee accurate translation of the construct, the forward primer included an ATG codon upstream of the hybridizing sequence. The cDNA for DN22- CB_1 was amplified using HF Platinum DNA polymerase (Invitrogen) and inserted into a PCRII-Topo vector (Invitrogen) according to the manufacturers' instructions. The absence of amplification mismatches was then verified by DNA sequencing. Primers used were: forward, with the inserted ATG in bold, 5'-ATGGTGGGCTCAAATGACATTCAG-3'; reverse, with the stop codon in bold, 5'-TCACAGAGCCTCGGCAGACGTG-3'. The cDNA sequence for CB_1 or DN22- CB_1 was inserted into a modified version of a pcDNA3.1 mammalian expression vector using BamHI–EcoRV according to standard cloning procedures. This modification allowed the co-expression of CB_1 or DN22- CB_1 with an mCherry fluorescent protein for control of transfection efficiency. For the study of mitochondrial motility, the coding sequence of CB_1 or DN22- CB_1 was fused to GFP using the pEGFP-N1 vector (Addgene) according to the manufacturer's instructions.

For MLS-PKA-CA, a forward primer including a restriction site after the initial ATG codon for future subcloning with mitochondrial leading sequences was coupled to a reverse primer hybridized to the end of the coding sequence of the catalytic subunit of PKA (pET15b PKA Cat from Addgene)³³, including a myc epitope and a TGA stop codon. Subsequently, the construct was subcloned into a pcDNA3.1 vector as an intermediate step and the QuikChange Multi Site-Directed Mutagenesis Kit (Agilent Genomics, Santa Clara, CA, USA) was used to mutate histidine-87 to glutamine, and tryptophan-196 to arginine to generate a constitutively active form of PKA²⁴. Finally, the construct was fused to a 4×MLS sequence to target the constitutively active PKA to mitochondria (MLS-PKA-CA). The absence of amplification mismatches and confirmation of mutagenesis was then verified by DNA sequencing. Primers used were: forward, with the inserted ATG in bold, 5'-TATCTGGATCCCTATGCAATTGGGCAACGCCG CCGCCGCCAAGAAGG-3'; reverse, with the stop codon and the myc epitope in bold, 5'-TATGATCTAGAGATCAGATCCTCTTCTGAGATGAGTTTT TGTTCAAACTCAGTAACTCCTTGCCCACTTC-3'; and for mutagenesis of H87, 5'-AAAGCAGATCGAGCAAACTCTGAATGAGAAG-3'; and W196 5'-GTGAAAGGCCGTACTAGGACCTTGTGTGGGA-3' (in bold are the mutated codons).

The phosphomimetic version of NDUFS2 was custom synthesized by Eurofins Genomics (Germany). Briefly, the NDUFS2 sequence (NM_153064) was modified

to obtain a phosphomimetic form mutating the 3 potential phosphorylation sites of PKA. The sites were chosen because consensus for their PKA phosphorylation nature was found between the two online available phosphorylation prediction algorithms, PhosphoMotif Finder (http://www.hprd.org/PhosphoMotif_finder) and PKA prediction site (<http://mendel.imp.ac.at/pkaPS/>)^{25,26}. By this approach, four sites were identified. One of these was excluded, because it is present on the mitochondrial leading sequence of NDUFS2. Thus, serines 296, 349 and 374 were mutated to aspartic acid to obtain a phosphomimetic version of NDUFS2 (NDUFS2-PM). A myc epitope was added at the C terminus of the protein for detection.

Preparation of adeno-associated viruses (AAV). The cDNAs coding for mouse CB_1 , DN22- CB_1 , MLS-PKA-CA, NDUFS2-PM and for GFP were subcloned into the pAM-CBA vector using standard molecular cloning techniques. The resulting vectors were transfected by calcium phosphate precipitation into HEK293 cells together with the rAAV-helper-plasmid pF6 and AAV1/2-serotype-packaging plasmids pRV1 and pH21 (ref. 34.). The viruses were then purified and titred as previously described^{35,36}. Virus titres were between 10^{10} and 10^{11} genomic copies per ml for all batches of virus used in the study.

Cell culture and transfection. All cell lines were originally obtained from ATCC (https://www.lgcstandards-atcc.org/Products/Cells_and_Microorganisms/Cell_Lines.aspx?geo_country=fr). Mouse 3T3 cells (3T3 F442A), HeLa and HEK293 cells were grown in Dulbecco modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 4.5 g l^{-1} glucose, 2 mM glutamine, 1 mM pyruvate. HEK293 cells were transfected with control plasmid, CB_1 or DN22- CB_1 cDNA coupled with mCherry cloned in pcDNA 3.1(+), respectively. Cells were transfected with sAC-HA or mtsAC-HA provided by G. Manfredi, (see refs 17, 21) or small hairpin RNA (shRNA) targeting AKAP121 provided by A. Feliciello (see ref. 22). HeLa cells were transfected with MLS-PKA-CA or NDUFS2-PM (see above). The transfections were carried out using EugeneHD (Roche, France) for 3T3 cells and polyethylenimine (PEI, Polysciences, USA) for HEK293 and HeLa cells, according to the manufacturers' protocols.

For biochemical experiments, primary hippocampal cultures were prepared from $CB_1^{+/+}$ and $CB_1^{-/-}$ P0–P1 mice. Briefly, after mice were killed by decapitation, hippocampi were extracted in dissection medium (10 mM HEPES, 0.3% glucose in Hank's balanced salt solution, pH 7.4) and dissociated in 0.25% trypsin for 30 min. Where indicated, dissociated cells were transfected with sAC-HA using the Amaxa P3 primary cell 4D-nucleofector kit (Lonza, France), according to the manufacturer's protocol. Cells were plated on poly-L-lysine-coated 96-well dishes using neurobasal/B27 medium (supplemented with 5% FBS, 2 mM L-glutamine, 1 mM pyruvate, 1 mM sodium lactate, 0.3% glucose and 37.5 mM NaCl) at a density of 50,000 cells per well. One hour after plating, the serum was removed. Primary hippocampal cultures contained both neurons and astrocytes, and were used at 3 days *in vitro* (DIV).

For live imaging of mitochondrial mobility, primary hippocampal cultures were prepared from $CB_1^{-/-}$ P0–P1 mice. Brains were extracted in PBS containing 0.6% glucose and 0.5% bovine serum albumin (BSA) and the hippocampi were dissected. To dissociate cells, a kit for dissociation of postnatal neurons was used following the manufacturer's instructions (Milteny Biotec, France). Cells were seeded onto 0.5 mg ml⁻¹ poly-L-lysine-coated 35-mm glass-bottom dishes (MatTek Corporation, France) for live imaging in neurobasal medium (Gibco, France) containing 2 mM L-glutamine, 120 µg ml⁻¹ penicillin, 200 µg ml⁻¹ streptomycin and B27 supplement (Invitrogen, France), and were maintained at 37°C in the presence of 5% CO₂. Cells were cultured for 7 to 9 days. Neuron transfection was carried out at 4–5 DIV, using a standard calcium phosphate transfection protocol, with a 1:2 DNA ratio of plasmids expressing pDsred2-mito³⁷ to GFP, CB_1 -GFP or DN22- CB_1 -GFP, respectively. Axonal mitochondrial mobility was recorded 72–96 h after transfection (see below). Cannabinoid treatments altered the percentage of axonal mobile mitochondria, without altering velocity, dwelling time or travelled distance (data not shown).

Mouse fibroblasts were generated from P0–P1 $CB_1^{-/-}$ pups. After mice were killed by decapitation, the dorsal skin was excised and minced in PBS. Cells were then separated by incubation in 0.25% trypsin solution in PBS, collected by centrifugation and resuspended in DMEM with 10% fetal bovine serum, 1% L-glutamine and 2% penicillin/streptomycin solution (Invitrogen, France). Cells were seeded in 25-cm² flasks and then expanded in 75-cm² flasks until reaching 90% confluence. Transfections were carried out by using a BTX-electroporator ECM 830 (Harvard Apparatus, France) (175 V, 1-ms pulse, five pulses, 0.5-s interval between pulses). Cells were electroporated in Optimum medium (Invitrogen, France) at 2×10^7 cells per ml (fibroblasts from two 75 cm² flasks at 90% confluence in 300 µl) in a 2-mm gap cuvette using 30 µg of either control plasmid (mCherry), CB_1 or DN22- CB_1 cDNA coupled with mCherry, respectively. After electroporation, cells were resuspended in DMEM with 10% fetal bovine serum, 1% L-glutamine and 2% penicillin/streptomycin solution (Invitrogen, France) and

seeded in three 100 cm² Petri dishes. All cells were maintained at 37°C and 5% CO₂ and collected 48 h after transfection for respiration experiments.

Isolation of mitochondria and cell fractionation. The brains of *CB₁^{+/-}* and *CB₁^{-/-}* littermates were dissected and mitochondria were purified using a Ficoll gradient as previously described^{7,8}. In brief, brains were extracted in ice-cold isolation buffer (250 mM sucrose, 10 mM Tris, 1 mM EDTA, pH 7.6) containing protease inhibitors (Roche, France) and 2 M NaF and homogenized with a Teflon potter. Homogenates were centrifuged at 1,500g for 5 min (4°C). The supernatant was then centrifuged at 12,500g (4°C). The pellet was collected and the cycle of centrifugation was repeated. To purify mitochondria, the final pellet was resuspended in 400 µl of isolation buffer, layered on top of a discontinuous Ficoll gradient (10% and 7% fractions) and centrifuged at 100,000g for 1 h (4°C). Purified mitochondria were recovered from the pellet obtained after ultracentrifugation. All experiments using freshly isolated brain mitochondria were performed within 3 h after purification.

The 3T3 cells were collected, resuspended in isolation buffer and disrupted with 25 strokes using a 25G needle. The total cell lysate was centrifuged at 500g (4°C) to remove cells debris and nuclei. The supernatant was kept and centrifuged at 12,500g for 10 min (4°C). The supernatant was then kept (cytosolic fraction), the pellet was resuspended, and the centrifugation cycle was repeated. Finally, the mitochondrial fractions were obtained from the last pellet.

Oxygen consumption measurements. The oxygen consumption of isolated mitochondria, homogenized hippocampus and cell lines was monitored at 37°C in a glass chamber equipped with a Clark oxygen electrode (Hansatech, UK). Purified mitochondria (75–100 µg) were suspended in 500 µl of respiration buffer (75 mM mannitol, 25 mM sucrose, 10 mM KCl, 10 mM Tris-HCl pH 7.4, 50 mM EDTA) in the chamber. Respiratory substrates were added directly to the chamber. Pyruvate (5 mM), malate (2 mM) and ADP (5 mM) were successively added to measure complex-I-dependent mitochondrial respiration. Complex-II-dependent respiration was measured using rotenone (0.5 µM), succinate (10 mM) and ADP (5 mM). Complex-IV-dependent respiration was measured using *N,N,N',N'*-tetramethyl-*p*-phenylenediamine (TMPD, 0.5 mM) and ascorbate (2 mM), in the presence of ADP (5 mM) and antimycin A (2.5 µM). Complex-I-dependent respiration was evaluated, unless stated otherwise.

For respiration in homogenized hippocampi, both hippocampi of each mouse were dissected and homogenized with a Teflon potter in 800 µl Mir05 buffer (Mitochondrial Physiology Network: 0.5 mM EGTA, 3 mM MgCl₂, 60 mM lactobionate, 20 mM taurine, 10 mM KH₂PO₄, 20 mM HEPES, 110 mM sucrose and 1 g l⁻¹ BSA) containing 12.5 µg ml⁻¹ of saponin. Subsequently, 15 µl of the homogenate was diluted in 1 ml Mir05 buffer and the oxygen consumption was measured with the respiratory substrates pyruvate (5 mM), malate (2 mM) and ADP (5 mM) to measure complex-I-dependent mitochondrial respiration before and after WIN (100 nM) or vehicle addition. Oligomycin (2 µg ml⁻¹), FCCP (0.5 µM), rotenone (0.5 µM) and antimycin A (2.5 µM) were injected subsequently into the chamber as modifiers of the respiration. 50 µl of the homogenate were saved for WB and protein quantification experiments.

The experiments using cell lines were performed on 2×10^6 cells ml⁻¹ in growth medium. Intact cells were transferred directly into the chamber and basal respiration was recorded. Drugs were added directly into the chambers. Mitochondria were incubated with PTX, KH7 and H89 for 5 min before addition of CB₁ agonists. HCO₃⁻ and 8-Br-cAMP were added 5 min after the addition of CB₁ agonists.

Oxygen consumption of primary hippocampal cultures was monitored using an XF96 Seahorse Bioscience analyser (Seahorse Bioscience, Denmark), according to the manufacturer's protocol. When indicated, oligomycin (2 µg ml⁻¹) and FCCP (1 µM) were injected directly into the wells. Other drugs were directly added into the medium 1 h before measurements.

Respiration of HEK293 cells co-expressing CB₁ and NDUFS2-PM or MLS-PKA-CA was analysed using the Oxygraph-2k (Oroboros Instruments, Austria). These experiments were performed on 5×10^5 cells ml⁻¹ in growth medium. WIN was directly added into the medium 30 min before measurements. Then, intact cells were transferred into the 2-ml chamber and basal respiration was recorded. **NADH-ubiquinone oxidoreductase (complex I) activity assay.** NADH oxidation into NAD⁺ by the first complex of the respiratory chain is coupled to the reduction of ubiquinone (coenzyme Q). The rate of this reaction is analysed by the measurement of NADH disappearance, which is spectrophotometrically detected (SAFAS, UVmc2) at 340 nm. The NADH extinction coefficient is 6.22 mM⁻¹ cm⁻¹. Final composition of the reaction solution was 50 mM K₂HPO₄ pH 7.2, 2.5 mg ml⁻¹ BSA, 0.1 mM ubiquinone and 200 µg total cell extract proteins or 50 µg purified brain mitochondria. The reaction was initiated by adding 0.1 mM NADH. The assay was monitored at 37°C for 5 min.

ATP content measurements. The intracellular ATP content was measured using the bioluminescent ATP kit HS II (Roche, France). *CB₁^{+/-}* and *CB₁^{-/-}* primary

hippocampal cultures (50,000 cells per well in a 96-well dish) were treated with THC (1 µM), WIN (1 µM) or vehicle in the presence or absence of rotenone (0.1 µM) for 1 h. Then, ATP measurements were performed as previously described³⁸. In brief, cells were lysed to release the intracellular ATP using the lysis buffer provided with the kit (equal volume) for 20 min. The lysate was then analysed in a 96-well plate luminometer (Luminoskan, Thermo Scientific, France) using the luciferine/luciferase reaction system provided with the kit. For this, 100 µl of luciferine/luciferase was injected in the wells and after 10 s of incubation, bioluminescence was read (1 s integration time). Standardizations were performed with known quantities of standard ATP provided with the kit. The ATP content derived from mitochondria was determined by subtracting ATP_{rotenone} values from the ATP_{total}: ATP_{mito} = ATP_{total} - ATP_{rotenone}.

Trypsin-sensitivity assays. 100 µg of mitochondria were suspended in isolation buffer, untreated or incubated with 0.01% trypsin in the presence or absence of 0.05% triton X-100 for 15 min at 37°C. Proteins were then processed for western immunoblot analyses.

Immunoprecipitation assays. Freshly purified brain mitochondria were resuspended in PBS (5 mg ml⁻¹) supplemented with protease inhibitor cocktail (Roche, France) and 2 mM NaF, and solubilized with 1% lauryl maltoside for 30 min (4°C). For co-immunoprecipitation of sAC and G_α, mitochondria were incubated with THC (800 nM) or vehicle for 5 min at 37°C. Proteins were incubated with a C-terminal anti-CB₁ antibody (Cayman, USA) or sAC R21 antibody (CEP Biotech, USA) overnight (4°C). For immunoprecipitation of complex I, mitochondrial proteins were treated with THC (800 nM), HCO₃⁻ (5 mM), 8-Br-cAMP (500 µM) or vehicle for 5 min at 37°C and then incubated with complex-I-agarose-conjugated beads (Abcam, UK). Protein A/G PLUS-agarose beads (Santa Cruz, USA) were then added and the incubation continued for 4 h (4°C). The elution was performed using glycine buffer (0.2 M glycine, 0.05% lauryl maltoside pH 2.5) and samples were processed for western immunoblotting.

ERK phosphorylation assays. Following transfection (mCherry, CB₁ or DN22-CB₁, respectively), cells were allowed to recover in serum containing medium for 24 h. Cells were then starved overnight in serum-free DMEM before treatment and lysis. The cells were then treated at 37°C with HU210 (100 nM) or vehicle for 10 min. The medium was rapidly aspirated and the samples were snap-frozen in liquid nitrogen and stored at -80°C before preparation for western blotting.

Protein extraction and western immunoblotting. For ERK-phosphorylation assays, lysis buffer (1 mM EGTA, 50 mM NaF, 1 mM Na₃VO₄, 50 mM Tris pH 7.5, 1% triton X-100, protease inhibitors, 30 mM 2-mercaptoethanol) was added and the cells were collected by scraping and pelleted by centrifugation at 12,500g (4°C) for 5 min to remove cell debris. Protein concentrations were measured using the Pierce BCA protein assay kit (Thermo Scientific), loaded with Laemmli buffer and kept at -80°C.

For western immunoblotting, the proteins were separated on Tris-glycine 7%, 10% or 12% acrylamide gels and transferred to PVDF membranes. Membranes were soaked in 5% milk (5% BSA for phosphorylation immunoblots) in tris-buffered saline (TBS; Tris 19.82 mM, NaCl 151 mM, pH 7.6) containing tween20 (0.05%). Mitochondrial proteins were immunodetected using antibodies against complex III core 2 (Abcam, ab14745; 1:1,000, 1 h, room temperature), succinate dehydrogenase subunit A (Abcam, ab14715; 1:10,000, 1 h, room temperature), NDUFA9 (Abcam, ab14713; 1:1,000, 1 h, room temperature), NDUFS2 (Abcam, ab110249; 1:1,000, 1 h, room temperature) and TOM20 (Santa Cruz, sz-11415; 1:1,000, 1 h, 4°C). Cytosolic proteins were probed with LDHa (Santa Cruz, sz-137243; 1:500, overnight, 4°C). Samples were also probed with antibodies against G_α proteins (Enzo Life Science, SA-126; 1:1,000, 1 h, room temperature), sAC (CEP Biotech, sAC R21; 1:500, overnight, 4°C), PKA (cAMP protein kinase catalytic subunit, Abcam, ab76238; 1:1,000, 1 h, room temperature), an antiserum directed against the C terminus of CB₁ receptor (Cayman, 10006590; 1:200, overnight, 4°C), AKAP121 (from A. Feliciello; 1:1,000, overnight, 4°C), PKA-dependent phosphorylation sites (phospho (Ser/Thr)-PKA substrate, Cell Signaling, 9621; 1:1,000, overnight, 4°C) and HA (Abcam, ab18181; 1:500, overnight, 4°C), p-ERK (phospho-p44/42 MAPK) corresponding to residues around Thr202/Tyr204 (Cell Signaling, 4370; 1:1,000, overnight, 4°C), ERK (p44/p42 MAPK; Cell Signaling, 9102; 1:2,000, 1 h, room temperature).

Mitochondrial proteins were also separated by two-dimension electrophoresis as described³⁹. Purified brain mitochondria were solubilized (10 mg ml⁻¹) in 0.75 M aminocaproic acid, 50 mM BisTris, (pH 7.0) with 1.5% *n*-dodecyl-maltoside for 30 min on ice, and were then centrifuged at 16,000g (4°C). The supernatant was collected and supplemented with 0.25% coomassie blue G and protease inhibitors (Roche, France). Proteins were then separated with 4–16% gradient native-PAGE gels (Invitrogen, France). The different lanes were cut out and processed for the second dimension on 12.5% SDS-PAGE gels after denaturation and reduction in 1% (w/v) sodium dodecyl sulphate and 1% (v/v) mercaptoethanol. The second

dimension gels were immunoblotted for detection of PKA-dependent phosphorylated proteins. A second-dimension gel was kept for coomassie blue staining. Then, membranes were washed and incubated with appropriate secondary horseradish peroxidase (HRP)-coupled antibodies (1 h, room temperature). Finally, the HRP signal was detected using the ECL-plus reagent (Amersham) and the Bio-Rad Quantity One system. Labelling was quantified by densitometric analysis using ImageJ (NIH) software.

Immunocytochemistry. HeLa cells were fixed in 4% formaldehyde dissolved in PBS (0.1 M, pH 7.4) and then washed with PBS. Cells were pre-incubated in a blocking solution of 10% normal goat serum, 0.1% triton X-100, 0.05% deoxycholate and 0.2 M glycine prepared in PBS for 1 h and then incubated with primary antibody rabbit anti-TOM20 (Santa Cruz, sc-11415; 1:500) and mouse anti-myc (Roche, 11667149001; 1:500) for 2 h in the same blocking solution. The cells were then washed in PBS for 1 h and were then incubated with fluorescent anti-mouse Alexa488 or anti-rabbit Alexa561 (Jackson ImmunoResearch; 1:800) in blocking solution for 1 h. Finally, cells were washed and mounted with fluoromont-G (Electron Microscopy Sciences). All the procedures were carried out at room temperature. The cells were analysed with a Confocal Leica DMI6000 microscope (Leica).

Proteomic analyses. Samples were digested by trypsin as previously described⁴⁰. Peptides were further analysed by nano-liquid chromatography coupled to a MS/MS LTQ-Orbitrap XL mass spectrometer (Thermo Fisher Scientific, Germany). Peptides were identified with SEQUEST and MASCOT algorithms through the Proteome Discoverer interface (Thermo Fisher Scientific, Germany) against a subset of the UniProt database restricted to Reference Proteome Set of *Mus musculus* (UniProtKB Release 2011_12, 14th December, 2011, 46,638 entries). Peptide validation was performed using Percolator algorithm⁴¹ and only 'high confidence' peptides were retained corresponding to a 1% false positive rate at peptide level.

PKA activity and cAMP content assays. Cyclic AMP levels and PKA activity of mitochondria isolated from the brain were assayed using the Direct Correlate-EIA cAMP kit (Assay Designs Inc., USA) and an ELISA kit (Enzo Life Science), respectively, according to the manufacturers' instructions. The different treatments described in the main text were performed at 37°C for 1 h.

Imaging and quantification of axonal mobility of mitochondria. Mitochondrial mobility in hippocampal neurons was recorded using an inverted Leica DMI6000 microscope (Leica Microsystems, Wetzlar, Germany) equipped with a confocal head Yokogawa CSU-X1 (Yokogawa Electric Corporation, Tokyo, Japan) and a sensitive Quantem camera (Photometrics, Tucson, USA). The diode lasers used were at 491 nm and 561 nm and the objective was a HCX PL APO CS 63 × oil 1.32 NA lens. The z stacks were obtained with a piezo P721.LLQ (Physik Instrumente (PI), Karlsruhe, Germany). The 37°C atmosphere during time-lapse image acquisition was created with an incubator box and an air heating system (Life Imaging Services, Basel, Switzerland) in the presence of 5% CO₂. This system was controlled by MetaMorph software (Molecular Devices, Sunnyvale, USA).

For mitochondrial axonal transport analysis, time-lapse series of image stacks composed of 6 images (512 × 512 pixels) were taken every 3 s for 15 min. HU210 was added just after the recording and 15 min later the same neuron was recorded for another 15 min. KH7 or vehicle were added 15 min before the first recording. All stacks obtained were processed first with MetaMorph software. Further image processing, analysis and video compilation (28 frames per second) and editing was done with ImageJ software (NIH, USA). Kymographs were generated with the KymoToolBox Plugin⁴². Between 10 and 32 axons were registered and analysed in each condition. In all cases, a mitochondrion was considered mobile when it moved more than 5 µm during the time of recording. Distances and speeds of retrograde and anterograde transport and dwelling time were measured separately from the corresponding kymographs, as previously described^{43,44}.

[³H]CP55,940 and [³⁵S]GTPγS binding studies. *Cell membrane microarray development.* The microarrays were composed of a collection of membrane homogenates isolated from HEK293 cells transfected with mCherry CB₁ or DN22-CB₁, or from hippocampi of CB₁^{+/+}(GFP), CB₁^{-/-}(GFP), CB₁^{-/-}(CB₁) or CB₁^{-/-}(DN22-CB₁) mice (see below), together with increasing amounts of BSA and membranes isolated from rat cerebral cortex, as positive internal controls⁴⁵. Briefly, samples were homogenized using a Teflon-glass grinder (Heidolph RZR 2020) or a disperser (Ultra-Turrax T10 basic, IKA) in 20 volumes of homogenized buffer (1 mM EGTA, 3 mM MgCl₂, and 50 mM Tris-HCl pH 7.4) supplemented with 250 mM sucrose. The crude homogenate was subjected to a 40g centrifugation for cells or 200g for tissue for 5 min, and the resultant supernatant was centrifuged again at 18,000g for 15 min (4°C, Microfuge 22R centrifuge, Beckman Coulter). The pellet was washed in 20 volumes of homogenized buffer and re-centrifuged under the same conditions. The homogenate aliquots were stored at -80°C until use. Protein concentration was measured by the Bradford method and adjusted

to the required concentrations. Microarrays were fabricated by a non-contact microarrayer (Nano_plotter NP 2.1) placing the cell membrane homogenates (4 nl per spot, 3–5 replicates per sample) onto glass slides⁴⁶. Microarrays were stored at -20°C until use.

[³H]CP55,940 binding assays. After thawing, cell membrane microarrays were incubated in assay buffer (50 mM Tris-Cl; 1% BSA; pH 7.4) for 30 min at room temperature. A second incubation was performed using the same buffer for 120 min at 37°C in the presence of [³H]CP55,940 (3 nM). Non-specific binding was determined with 10 µM WIN55,212-2. Afterwards, microarrays were washed twice in buffer, dipped in deionized water and dried. Finally, they were exposed to films, developed, scanned and quantified as described below.

[³⁵S]GTPγS binding studies. [³⁵S]GTPγS binding studies were carried out according to the patented methodology for the screening of molecules that act through G-protein-coupled receptors using cell membrane microarrays⁴⁵. Briefly, thawed cell membrane microarrays were dried 20 min at room temperature and were subsequently incubated in assay buffer (50 mM Tris-Cl; 1 mM EGTA; 3 mM MgCl₂; 100 mM NaCl; 0.5% BSA; pH 7.4) for 15 min at room temperature. Microarrays were transferred into assay buffer containing 50 µM GDP and 0.1 nM [³⁵S]GTPγS, with the cannabinoid agonists WIN55,212-2 or HU210, at increasing concentrations, and incubated at 30°C for 30 min. Non-specific binding was determined with GTPγS (10 µM). After washing, microarrays, together with ARC [¹⁴C]-standards, were exposed to films, developed, scanned and quantified. The protein concentration in each spot was measured by the Bradford method and used to normalize the [³⁵S]GTPγS binding results to nCi per ng protein. Data from the dose-response curves (5 replicates in triplicate) were analysed using the program Prism (GraphPad Software Inc., San Diego, CA) to yield EC₅₀ (effective concentration 50%) and E_{max} (maximal effect) of the drugs on each different sample by nonlinear regression analysis. Samples displaying [³H]CP55,940 binding below the values of hippocampi from CB₁^{-/-} mice were excluded from [³⁵S]GTPγS binding analysis.

Surgery and drug administration. Mice (7–9 weeks of age) were anaesthetized by i.p. injection of a mixture of ketamine (100 mg kg⁻¹; Imalgene 500, Merial) and Xylazine (10 mg kg⁻¹; Rompun, Bayer) and placed into a stereotaxic apparatus (David Kopf Instruments) with mouse adaptor and lateral ear bars.

For intracerebroventricular injections of drugs, mice were unilaterally implanted with a 1.0-mm stainless-steel guide cannula targeting the lateral ventricle with the following coordinates: anterior-posterior -0.2; lateral ± 0.9; dorsal-ventral -2.0. For intrahippocampal injections of drugs, mice were bilaterally implanted with 1.0-mm stainless-steel guide cannulae targeting the hippocampus with the following coordinates: anterior-posterior -3.1; medial-lateral ± 1.3; dorsal-ventral -0.5. Guide cannulae were secured with cement anchored to the skull by screws. Mice were allowed to recover for at least one week in individual cages before the start of experiments. Mice were weighed daily and individuals that failed to return to their pre-surgery body weight were excluded from subsequent experiments. The intrahippocampal and intracerebroventricular drug injections were performed by using injectors protruding 1 mm from the tip of the cannula.

In vivo viral re-expression of CB₁ and DN22-CB₁. For viral intrahippocampal AAV delivery, mice were submitted to stereotaxic surgery (as above) and AAV vectors were injected with the help of a microsyringe (0.25-ml Hamilton syringe with a 30-gauge bevelled needle) attached to a pump (UMP3-1, World Precision Instruments). Mice were injected directly into the hippocampus (0.5 µl per injection site at a rate of 0.5 µl per min), with the following coordinates: dorsal hippocampus, anterior-posterior -1.8; medial-lateral ± 1; dorsal-ventral -2.0 and -1.5; ventral hippocampus: anterior-posterior -3.5; medial-lateral ± 2.7; dorsal-ventral -4 and -3. Following virus delivery, the syringe was left in place for 1 min before being slowly withdrawn from the brain. CB₁^{+/+} mice were injected with AAV-GFP to generate CB₁^{+/+}(GFP) mice; CB₁^{-/-} mice were injected with AAV-GFP, AAV-CB₁ or AAV-DN22-CB₁, to obtain CB₁^{-/-}(GFP), CB₁^{-/-}(CB₁) and CB₁^{-/-}(DN22-CB₁) mice, respectively. Animals were used for experiments 4–5 weeks after injections. Mice were weighed daily and individuals that failed to return to their pre-surgery body weight were excluded from subsequent experiments. CB₁-receptor expression was verified by fluorescent or electromicroscopic immunohistochemistry (see below).

In vivo viral modulation of mitochondrial PKA and NDUF52 activities. The AAV vectors, MLS-PKA-CA or NDUF52-PM were injected directly into the dorsal hippocampus (1.0 µl per injection site at a rate of 0.5 µl per min) of C57BL/6N mice, with the following coordinates: anterior-posterior -1.8; medial-lateral ± 1; dorsal-ventral -2.0 and -1.5. Following virus delivery, the syringe was left in place for 1–2 min before being slowly withdrawn from the brain. Animals were used for experiments 4–5 weeks after viral delivery. Mice were habituated to i.p. injections (saline) before the behavioural paradigm (see below). The hippocampal expression

of myc-tagged MLS-PKA-CA and NDUFS2-PM was verified by immunohistochemistry using anti-myc antibodies.

Fluorescence immunohistochemistry. Mice were anaesthetized with chloral hydrate (400 mg kg⁻¹ body weight), transcardially perfused with Ringer solution (NaCl (135 mM), KCl (5.4 mM), MgCl₂·6H₂O (1 mM), CaCl₂·2H₂O (1.8 mM), HEPES (5 mM)). Heparin choay (25,000 UI per 5 ml) was added extemporarily and tissues were then fixed with 500 ml of 4% formaldehyde dissolved in PBS (0.1 M, pH 7.4) and prepared at 4°C. After perfusion, the brains were removed and incubated several additional hours in the same fixative. Serial vibrosections were cut at 40–50-μm thickness and collected in PBS at room temperature. Sections were pre-incubated in a blocking solution of 10% donkey serum, 0.1% sodium azide and 0.3% triton X-100 prepared in PBS for 30 min–1 h at room temperature. Free-floating sections were incubated for 48 h (4°C) with goat anti-CB₁ polyclonal antibodies raised against a C-terminal sequence of 31 amino acids (NM007726) of the mouse CB₁ receptor (CB₁-Go-Af450-1; 2 μg ml⁻¹; Frontier Science Co. Ltd) or overnight (4°C) with rabbit anti-myc (Ozyme; 1:1,000). The antibody was prepared in 10% donkey serum in PBS containing 0.1% sodium azide and 0.5% triton X-100. Then, the sections were washed in PBS for 30 min at room temperature. The tissue was subsequently incubated with fluorescent anti-goat Alexa488 (1:200; Jackson ImmunoResearch) for 4 h and washed in PBS at room temperature, before being incubated with DAPI (1:20,000) for 10 min for nuclear counterstaining. Finally, sections were washed, mounted, dried and a coverslip was added on top with DPX (Fluka Chemie AG). The slides were analysed with an epifluorescence Leica DM6000 microscope (Leica).

Immuno-electron microscopy. CB₁^{+/+}(GFP), CB₁^{-/-}(GFP), CB₁^{-/-}(CB₁) and CB₁^{-/-}(DN22-CB₁) mice (*n* = 3 per group) were processed for electron microscope pre-embedding immunogold labelling as previously described^{7,8}. Immunodetection was performed in 50-μm-thick sections of hippocampus with goat anti-CB₁ polyclonal antibodies raised against a 31 amino acid C-terminal sequence (NM007726) of the mouse CB₁ receptor (Frontier Institute Co. Ltd, CB₁-Go-Af450-1; 2 μg ml⁻¹).

Semiquantification of mtCB₁. Immunogold particles were identified and counted. To exclude the risk of counting possible false positive mitochondrial labelling, we used strict semi-quantification methods of mtCB₁ receptors as recently described, excluding immunogold particles that were located on mitochondrial membranes but at a distance ≤80 nm from other cellular structures⁸. The normalized number of immunogold particles located on mitochondria versus the total amount of immunogold particles in each field was used to calculate the proportion of mtCB₁ receptors over total CB₁.

Field excitatory post-synaptic potentials (fEPSP) recording. Mice were anaesthetized with isoflurane and killed by decapitation. Brains were rapidly removed and chilled in an ice-cold, carbonated (bubbled with 95% O₂–5% CO₂) cutting solution containing 180 mM sucrose, 2.5 mM KCl, 0.2 mM CaCl₂, 12 mM MgCl₂, 1.25 mM NaH₂PO₄, 26 mM NaHCO₃ and 11 mM glucose (pH 7.4). Sagittal hippocampal slices (350-μm thick) were cut using a Leica VT1200S vibratome and incubated with artificial cerebrospinal fluid (ACSF) containing 123 mM NaCl, 1.25 mM NaH₂PO₄, 11 mM glucose, 2.5 mM KCl, 2.5 mM CaCl₂, 1.3 mM MgCl₂ and 26 mM NaHCO₃ (osmolality of 298 ± 7; pH 7.4) for 30 min at 34°C. The slices were subsequently transferred to a holding chamber, where they were maintained at room temperature until experiments. Slices were individually transferred to a submerged chamber for recording and continuously perfused with oxygenated (95% O₂–5% CO₂) ACSF (3–5 ml min⁻¹). All experiments were performed at room temperature. fEPSPs were recorded using glass micropipettes (2–4 mΩ) filled with normal ACSF positioned in the CA1 hippocampal region. Slices from the middle hippocampus were used preferentially. fEPSPs responses were evoked by stimulation (0.1-ms duration, 10–30-V amplitude) delivered to the stratum radiatum to stimulate the Schaffer collateral fibres using similar glass electrodes used for the recordings, in the presence of picrotoxin 100 μM. Recordings were obtained using an Axon Multiclamp 700B amplifier (Molecular Devices). Signals were filtered at 2 kHz, digitized, sampled and analysed using Axon Clampfit software (Molecular Devices). In CB₁^{-/-}(CB₁) and CB₁^{-/-}(DN22-CB₁), two slices (1 each) were excluded from analysis, because immunohistochemistry showed no re-expression.

Novel object-recognition memory task. To study the effect of mtCB₁ receptor signalling on cannabinoid-induced amnesia, we used the hippocampal-dependent NOR memory task in an L-maze (L-M/NOR)^{14,47,48}. As compared to other hippocampal-dependent memory tasks, this test presents several advantages for the aims of the present study: (i) the acquisition of L-M/NOR occurs in one step and previous studies revealed that the consolidation of this type of memory is deeply altered by acute immediate post-training administration of cannabinoids via hippocampal CB₁ receptors^{14,48}; (ii) this test allows repeated independent measurements of memory performance in individual animals⁴⁷, thereby allowing within-subject comparisons, eventually excluding potential individual differences

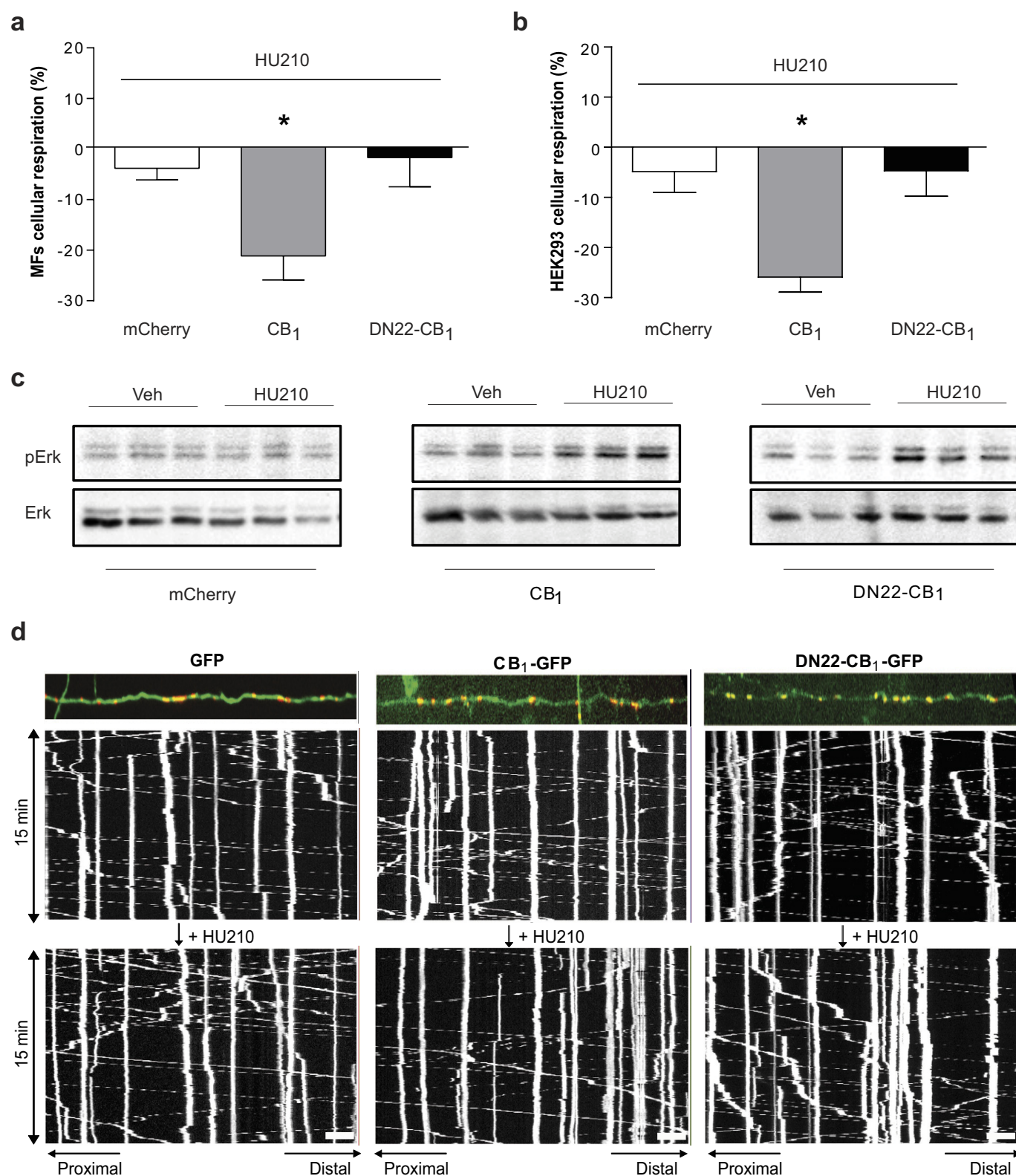
in viral infection and/or expression of proteins; (iii) notably, CB₁^{-/-} mice do not respond to the administration of cannabinoids, but they do not show any spontaneous impairment of performance in L-M/NOR¹⁴, thereby allowing the use of re-expression approaches to study the role of hippocampal mtCB₁ receptors in the cannabinoid-induced blockade of memory consolidation.

This task was performed with an L-maze made out of dark-grey Plexiglas with two corridors (35 cm and 30 cm long, respectively, for external and internal V walls, 4.5 cm wide and 15-cm high walls) set at a 90° angle and under a weak light intensity (50 Lux). The task consisted of 3 sequential daily trials of 9 min. Day 1 (habituation): mice were placed at the intersection of the two arms and were let free to explore the maze. Day 2 (acquisition): two identical objects were placed at the end of each arm. After 9 min of exploration, mice were removed and injected. Day 3 (retrieval): A novel object different in its shape, colour and texture was placed at the end of one of the arm, whereas the familiar object remained at the end of the other arm. The position of the novel object and the pairings of novel and familiar objects were randomized. Exploration of each object was scored off-line by at least two experienced observers blind to treatments and/or genotypes. Exploration was defined as the time spent by the mouse with the nose pointing to the object at a distance of less than 1 cm, whereas climbing on or chewing the object was not considered as exploration¹⁴. Memory performance was assessed by the discrimination index. The discrimination index was calculated as the difference between the time spent exploring the novel (TN) and the familiar object (TF) divided by the total exploration time (TN+TF): discrimination index = (TN–TF)/(TN+TF). Mice receiving the acute intrahippocampal infusion of KH7 (10 mM) and WIN (5 mg kg⁻¹ i.p., and mice that received intrahippocampal injection with AAV-MLS-PKA-CA or AAV-NDUFS2-PM were submitted to a single L-M/NOR session. Due to the limited numbers of available mice, null CB₁^{-/-} mice virally injected with AAV-GFP, AAV-CB₁ or AAV-DN22-CB₁ were tested twice with a one-week interval using different pairs of objects and treated the first time with vehicle and the following with WIN. Every pair of objects was previously screened to exclude that the animals might exhibit significant preference for any specific item. After the NOR task, the hippocampi of vehicle-treated AAV-MLS-PKA-CA and AAV-NDUFS2-PM animals were dissected and used for respiration experiments (see above). Expression of the myc epitope was verified by immunohistochemistry in the hippocampi of animals treated with WIN.

Statistical analyses. All graphs and statistical analyses were performed using GraphPad software (version 5.0 or 6.0). Results were expressed as means of independent data points ± s.e.m. For biochemical quantifications (cAMP levels, PKA and complex-I activities and oxygen consumption), data are presented as percentage of controls with or without the application of cannabinoid drugs. With the exception of KH7 (see Extended Data Fig. 5a), no other drugs and plasmids had any effect *per se* on any measured parameter (not shown). Data were analysed using paired or unpaired Student's *t*-test, one-way (followed by Tukey's *post hoc* test) or two-way ANOVA (followed by Bonferroni's *post hoc* test), as appropriate. Detailed statistical data for each experiment are reported in Supplementary Tables 1–3.

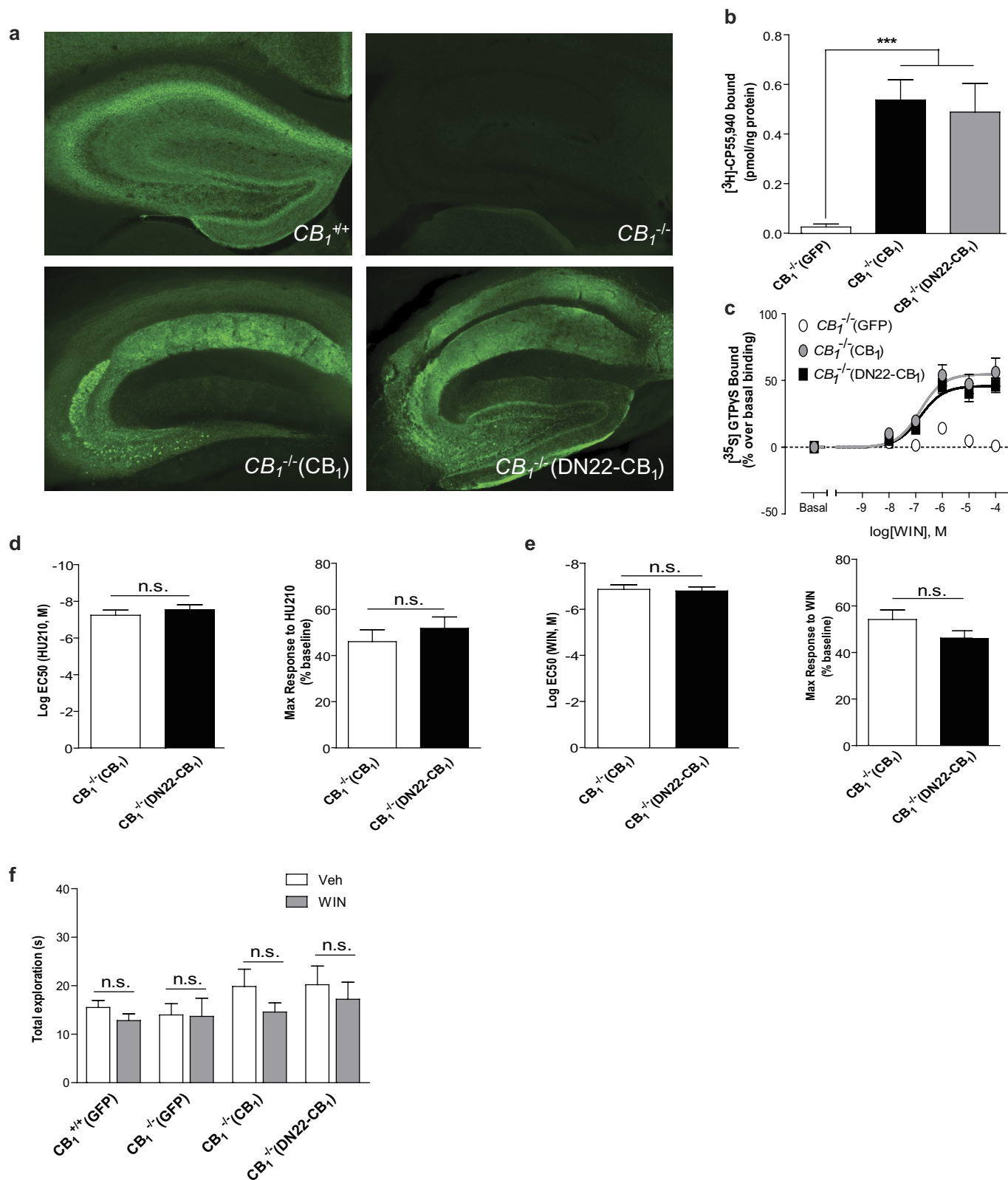
- Marsicano, G. *et al.* The endogenous cannabinoid system controls extinction of aversive memories. *Nature* **418**, 530–534 (2002).
- Martín-Couce, L. *et al.* Chemical probes for the recognition of cannabinoid receptors in native systems. *Angew. Chem. Int. Ed. Engl.* **51**, 6896–6899 (2012).
- Narayana, N., Cox, S., Shaltiel, S., Taylor, S. S. & Xuong, N. Crystal structure of a polyhistidine-tagged recombinant catalytic subunit of cAMP-dependent protein kinase complexed with the peptide inhibitor PKI(5–24) and adenosine. *Biochemistry* **36**, 4438–4448 (1997).
- McClure, C., Cole, K. L., Wulff, P., Klugmann, M. & Murray, A. J. Production and titration of recombinant adeno-associated viral vectors. *J. Vis. Exp.* **27**, e3348 (2011).
- Soria-Gómez, E. *et al.* The endocannabinoid system controls food intake via olfactory processes. *Nat. Neurosci.* **17**, 407–415 (2014).
- Chiarlone, A. *et al.* A restricted population of CB₁ cannabinoid receptors with neuroprotective activity. *Proc. Natl Acad. Sci. USA* **111**, 8257–8262 (2014).
- López-Doménech, G. *et al.* The Eutherian *Armxc* genes regulate mitochondrial trafficking in neurons and interact with Miro and Trak2. *Nat. Commun.* **3**, 814 (2012).
- Jose, C. *et al.* AICAR inhibits cancer cell growth and triggers cell-type distinct effects on OXPHOS biogenesis, oxidative stress and Akt activation. *Biochim. Biophys. Acta* **1807**, 707–718 (2011).
- Hébert Chatelain, E., Dupuy, J. W., Letellier, T. & Dachary-Prigent, J. Functional impact of PTP1B-mediated Src regulation on oxidative phosphorylation in rat brain mitochondria. *Cell. Mol. Life Sci.* **68**, 2603–2613 (2011).
- Pocaly, M. *et al.* Proteomic analysis of an imatinib-resistant K562 cell line highlights opposing roles of heat shock cognate 70 and heat shock 70 proteins in resistance. *Proteomics* **8**, 2394–2406 (2008).

41. Käll, L., Canterbury, J. D., Weston, J., Noble, W. S. & MacCoss, M. J. Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nat. Methods* **4**, 923–925 (2007).
42. Zala, D. *et al.* Vesicular glycolysis provides on-board energy for fast axonal transport. *Cell* **152**, 479–491 (2013).
43. De Vos, K. J. & Sheetz, M. P. Visualization and quantification of mitochondrial dynamics in living animal cells. *Methods Cell Biol.* **80**, 627–682 (2007).
44. Courchet, J. *et al.* Terminal axon branching is regulated by the LKB1-NUAK1 kinase pathway via presynaptic mitochondrial capture. *Cell* **153**, 1510–1525 (2013).
45. Rodríguez-Puertas, R., Barreda-Gómez, G., Giralt, M. & Fernández-Pastor, B. Method of quantifying the G protein-coupled receptor (GPCR)/G protein coupling using a cell membrane array. Patent US 8017346 B2. *Google Patents* (2013).
46. Rodríguez-Puertas, R. *et al.* Method for the surface treatment of solid substrates. Patent EP 2048534 A1 *Google Patents* (2009).
47. Busquets-Garcia, A. *et al.* Differential role of anandamide and 2-arachidonoylglycerol in memory and anxiety-like responses. *Biol. Psychiatry* **70**, 479–486 (2011).
48. Puighermanal, E. *et al.* Dissociation of the pharmacological effects of THC by mTOR blockade. *Neuropsychopharmacology* **38**, 1334–1343 (2013).



Extended Data Figure 1 | The mutant DN22-CB₁ receptor is functional, but it does not mediate cannabinoid-induced alterations of mitochondrial activity and mobility. **a**, Acute treatment with HU210 (100 nM) induces changes in cellular oxygen consumption in intact primary fibroblasts (MFs) from CB₁^{-/-} mice electroporated with a CB₁-expressing plasmid, whereas no alterations were observed in DN22-CB₁-expressing or mock-electroporated (mCherry) CB₁^{-/-} mouse fibroblasts ($n = 6-10$). **b**, HU210 (200 nM) treatment decreases HEK293 cellular respiration in cells expressing CB₁, but not mCherry or DN22-CB₁ receptors ($n = 4$). **c**, Representative immunoblotting showing the effect of

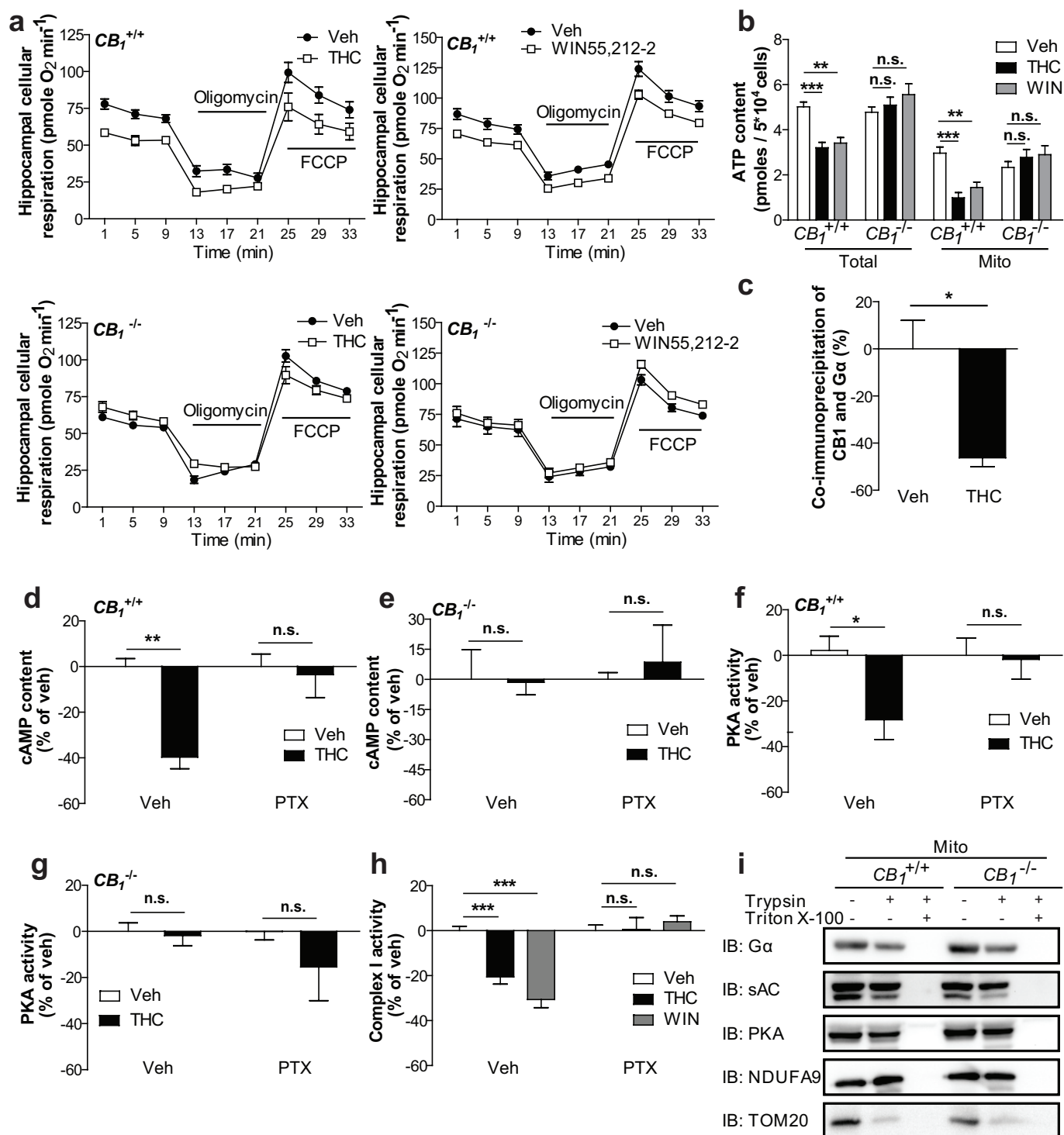
HU210 (100 nM) on pERK and ERK levels in HEK293 cells transfected with plasmids expressing mCherry, CB₁ or DN22-CB₁ ($n = 7-13$). Quantitative data are shown in Fig. 1b of main text. **d**, Representative confocal images (green, GFP; red, pDsred2-mito) and kymographs corresponding to the results shown in Fig. 1d; showing how HU210 (1 μ M) reduces mitochondrial mobility in the primary neurons from CB₁^{-/-} mice transfected with vectors expressing CB₁, but not GFP or DN22-CB₁. Scale bar, 10 μ M. Data, mean \pm s.e.m.; * $P < 0.05$. For statistics, see Supplementary Tables 1-3.



Extended Data Figure 2 | Characterization of *CB₁^{+/+}*(GFP), *CB₁^{-/-}*(GFP), *CB₁^{-/-}*(CB₁) and *CB₁^{-/-}*(DN22-CB₁) mice.

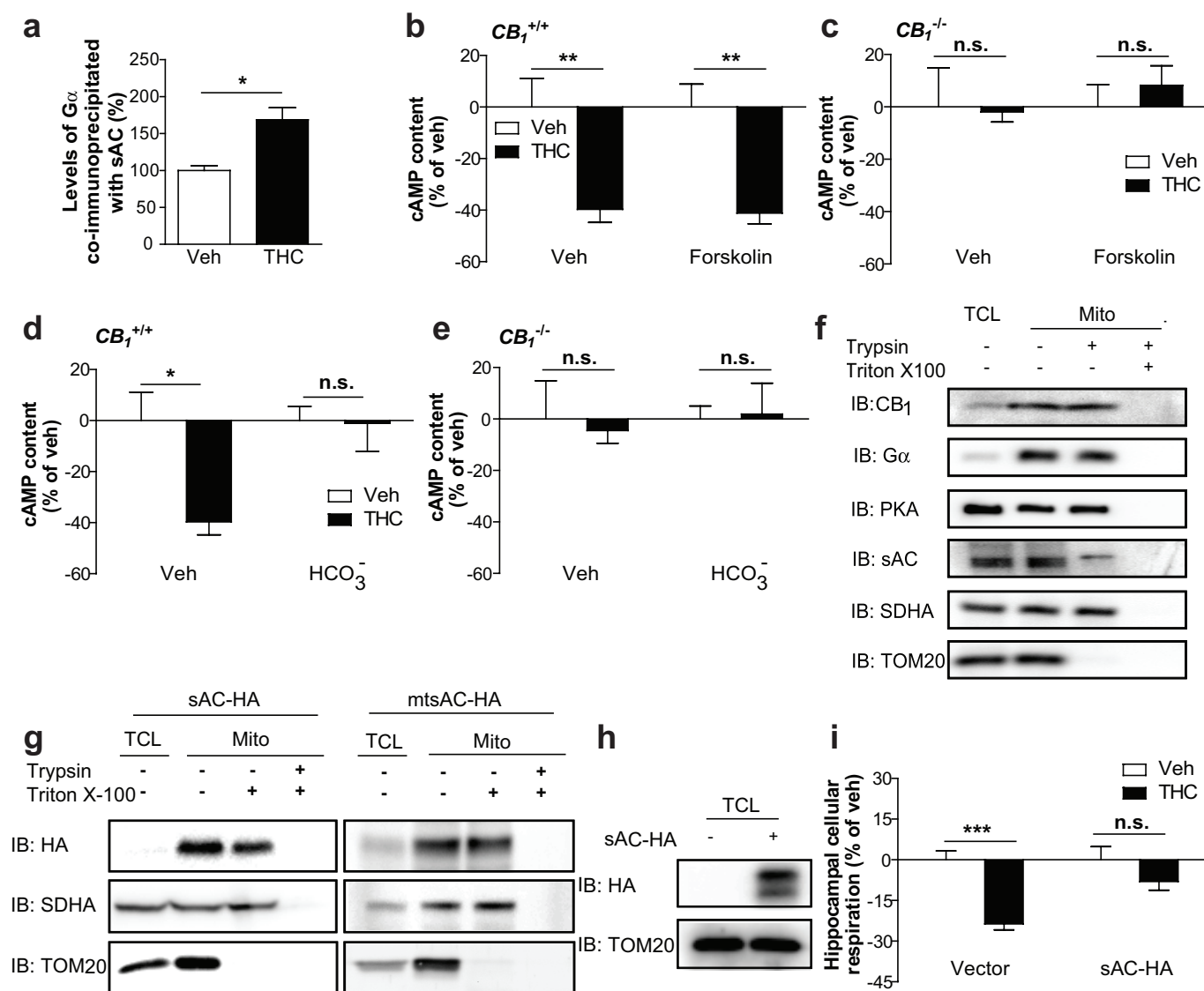
a, Representative immunofluorescence micrographs showing CB₁-receptor expression in hippocampal brain sections from *CB₁^{+/+}*(GFP), *CB₁^{-/-}*(GFP), *CB₁^{-/-}*(CB₁) and *CB₁^{-/-}*(DN22-CB₁) mice, respectively. **b**, [³H]CP55,940 CB₁-receptor binding experiments, showing similar levels of CB₁ protein expression in hippocampal extracts from *CB₁^{-/-}*(CB₁) and *CB₁^{-/-}*(DN22-CB₁) mice, both significantly higher than control *CB₁^{-/-}*(GFP) (*n* = 7–24). **c**, Dose-response curves of

[³⁵S]GTPγS binding in membranes isolated from hippocampi of *CB₁^{-/-}*(GFP), *CB₁^{-/-}*(CB₁) and *CB₁^{-/-}*(DN22-CB₁) mice treated with WIN (*n* = 9–24). **d**, **e**, CB₁-receptor-dependent activation of G proteins is similar in hippocampi from *CB₁^{-/-}*(CB₁) and *CB₁^{-/-}*(DN22-CB₁) mice, as shown by EC₅₀ dose and maximal activation values calculated from [³⁵S]GTPγS binding experiments after treatment with HU210 (**d**) or WIN (**e**). **f**, Total exploration time in NOR experiments as described in Fig. 1i. Data, mean ± s.e.m.; ****P* < 0.001; NS, not significant. For statistics, see Supplementary Tables 1–3.



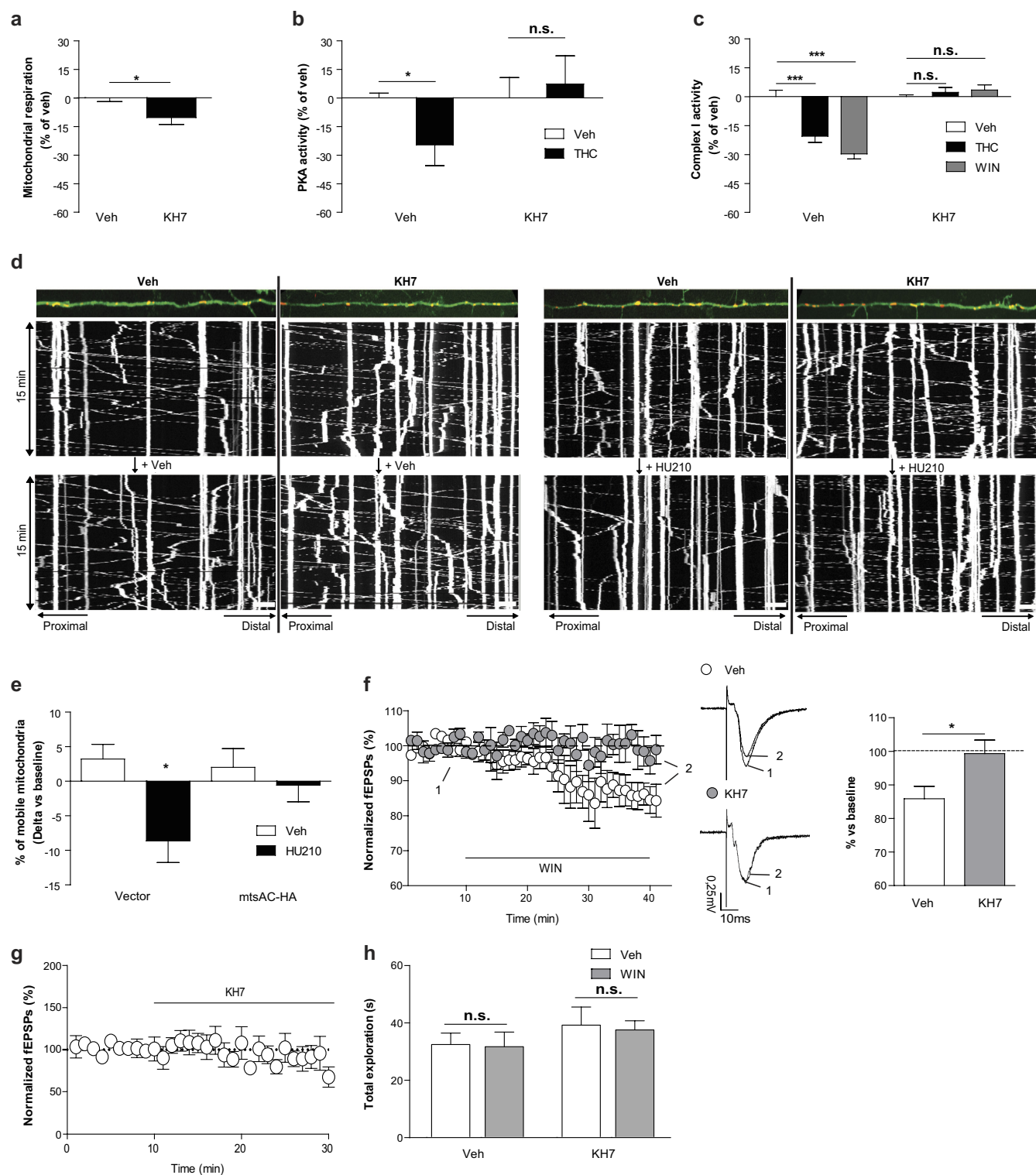
Extended Data Figure 3 | Activation of mtCB₁ receptor affects mitochondrial activity via G_{αi/o} signalling. **a**, THC (1 μM) or WIN (1 μM) decrease cellular respiration of primary hippocampal cultures derived from *CB1*^{+/+} but not from *CB1*^{-/-} mice (*n* = 5–15). **b**, THC (1 μM) or WIN (1 μM) reduce both total and mitochondrial ATP levels of primary hippocampal cultures derived from *CB1*^{+/+} but not from *CB1*^{-/-} mice (*n* = 6–14). **c**, Quantification of data presented in Fig. 2b, showing the release of the G protein upon activation of mtCB₁ receptors (*n* = 3). **d**, **e**, Effects of pertussis toxin (PTX, 1 μg ml⁻¹) and THC (800 nM) on cAMP levels in purified brain mitochondria obtained from *CB1*^{+/+} mice (**d**) and *CB1*^{-/-} littermates (**e**) (*n* = 4). **f**, **g**, Effects of PTX (1 μg ml⁻¹) and

THC (800 nM) on PKA activity in purified brain mitochondria obtained from *CB1*^{+/+} mice (**f**) and *CB1*^{-/-} littermates (**g**) (*n* = 5–6). **h**, PTX (1 μg ml⁻¹) blocks the effect of THC (800 nM) and WIN (100 nM) on complex-I activity of purified brain mitochondria obtained from *CB1*^{+/+} mice (*n* = 3–6). **i**, Representative trypsin-sensitivity assay (3 independent experiments) showing the intra-mitochondrial localization of G_α, sAC and PKA. NDUF9, a subunit of complex I; TOM20, translocase of outer membrane subunit 20. Data, mean ± s.e.m.; **P* < 0.05; ***P* < 0.01; ****P* < 0.001; NS, not significant. For statistics, see Supplementary Tables 1–3.



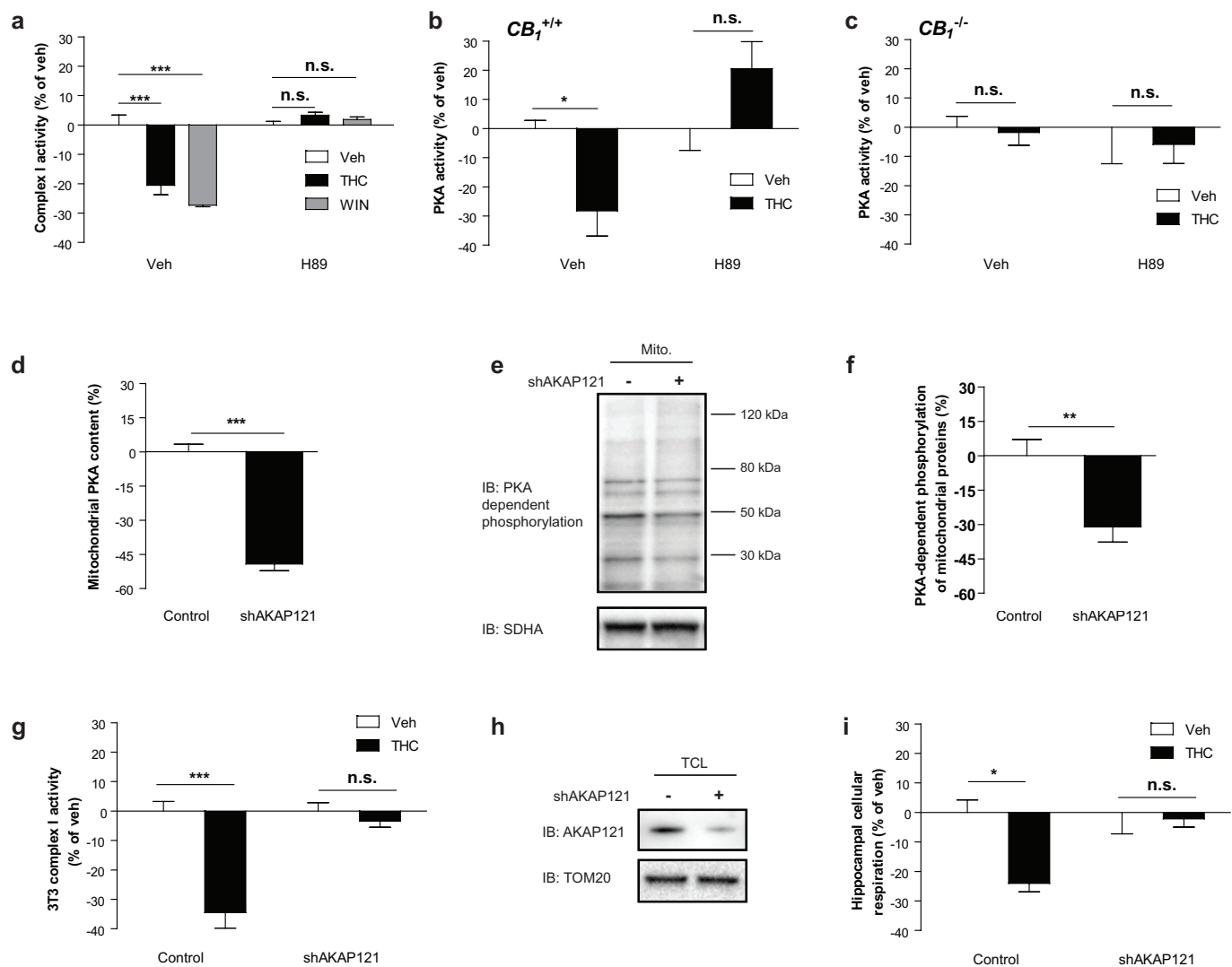
Extended Data Figure 4 | sAC mediates the mitochondrial effects of mtCB₁-receptor activation. **a**, Quantification of data presented in Fig. 2e, indicating higher interaction between G_{α} proteins and sAC after mtCB₁ activation ($n=3$). **b**, The transmembrane adenylyl cyclase activator forskolin (10 μ M) does not alter basal levels and does not reverse the effect of THC (800 nM) on cAMP content of $CB_1^{+/+}$ brain mitochondria ($n=4$). **c**, Forskolin (10 μ M) has no effect on basal cAMP levels of $CB_1^{-/-}$ brain mitochondria ($n=3-4$). **d**, The sAC activator HCO_3^- (5 mM) reverses the effect of THC (800 nM) on cAMP content in purified brain mitochondria from $CB_1^{+/+}$ mice ($n=4$). **e**, HCO_3^- (5 mM) and THC (800 nM) have no effect on cAMP levels ($n=3-4$) of $CB_1^{-/-}$ brain mitochondria. **f**, Representative trypsin sensitivity assay in mitochondrial fractions of

3T3 cells (3 independent experiments) showing intra-mitochondrial localization of CB₁ receptors, G_{α} proteins, AKAP121, PKA and sAC in 3T3 cells. SDHA, succinate dehydrogenase A (subunit of complex II); TOM20, translocase of outer membrane subunit 20. **g**, Representative trypsin-sensitivity assay (3 independent experiments) showing that sAC-HA and mtsAC-HA are localized inside mitochondria in transfected 3T3 cells. **h**, Representative immunoblot (3 independent experiments) showing the expression of sAC-HA in transfected and non-transfected primary hippocampal cultures. **i**, Expression of sAC-HA in primary hippocampal cultures blocks the effect of THC (1 μ M) on basal cellular respiration ($n=8-10$). Data, mean \pm s.e.m.; $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; NS, not significant. For statistics, see Supplementary Tables 1–3.



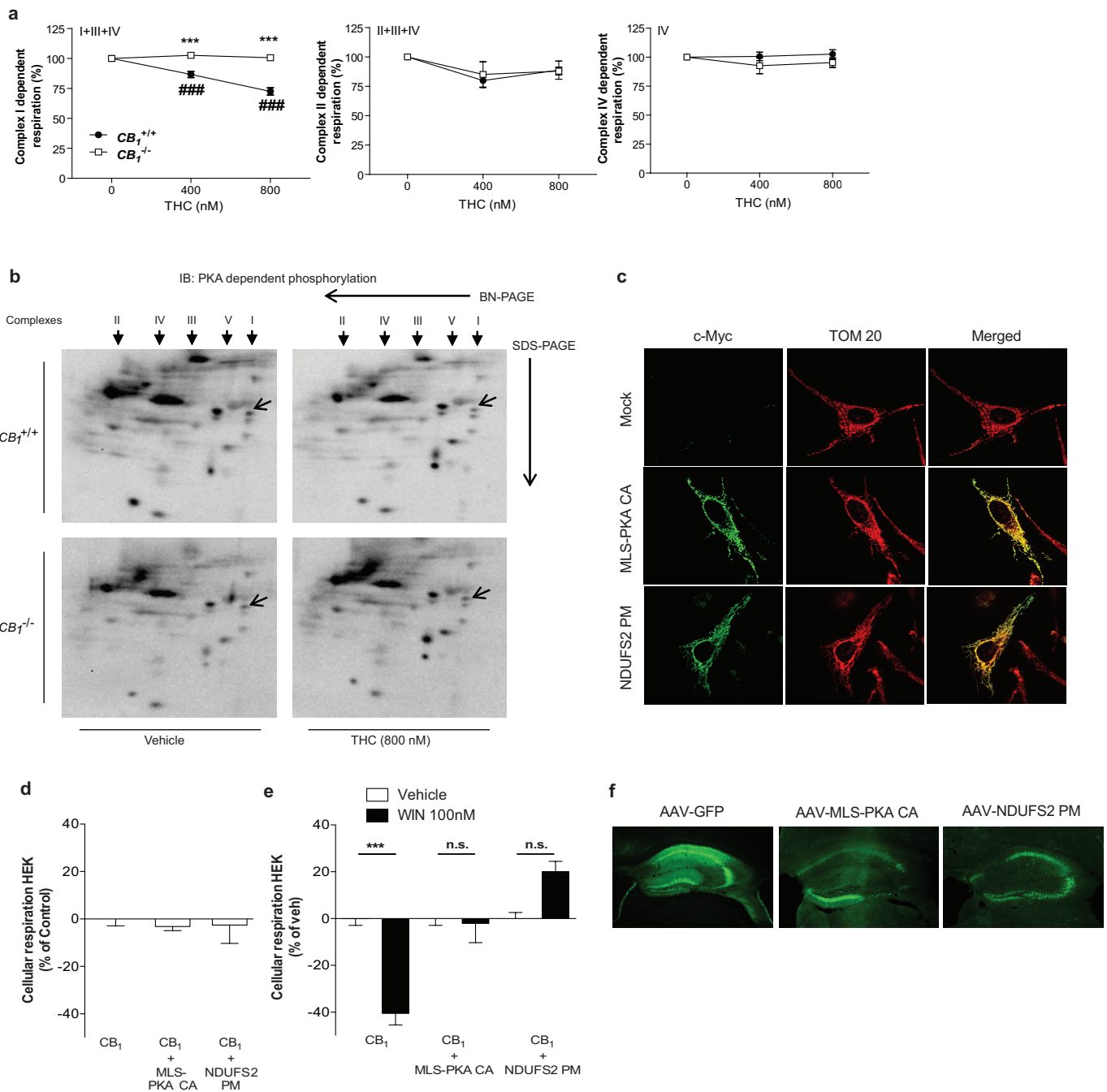
Extended Data Figure 5 | The sAC inhibitor KH7 prevents the effects of mtCB₁ activation. **a**, The effect of KH7 (5 μ M) on mitochondrial respiration in purified brain mitochondria obtained from CB₁^{+/+} mice ($n=5-6$). **b**, KH7 (5 μ M) blocks the effect of THC (800 nM) on PKA activity in purified brain mitochondria obtained from CB₁^{+/+} mice ($n=5-11$). **c**, KH7 (5 μ M) blocks the effect of THC (800 nM) and WIN (100 nM) on complex-I activity in purified brain mitochondria obtained from CB₁^{+/+} mice ($n=3-6$). **d**, Representative confocal images (green, CB₁-GFP; red, pDsred2-mito) of analysed axonal tracts and kymographs of the results shown in Fig. 3b, representing the effect of KH7 (5 μ M) on the reduction in mitochondrial mobility in primary hippocampal neurons induced by HU210 (1 μ M). Scale bar, 10 μ m. **e**, mtsAC-HA blocks the decrease in mitochondrial mobility in CB₁-expressing neurons induced by

HU210 (1 μ M) ($n=7-18$). **f**, The effect of KH7 (10 μ M) on the decrease in fEPSP at CA3-CA1 synapses from C57BL/6N mice induced by WIN (5 μ M). Left, plots of normalized fEPSP slopes *in vitro* with representative fEPSP traces before (1) and after (2) WIN incubation. Vehicle ($n=6$) or KH7 ($n=5$) were pre-incubated 10 min before cannabinoid application. Right, histogram summarizing the average changes in the percentage of the fEPSP slope before (100% baseline, dotted line) and after WIN treatment in the presence or absence of KH7. KH7 (10 μ M) alone does not affect the fEPSP slope at CA3-CA1 synapses from C57BL/6N mice. Plots of normalized fEPSP slopes *in vitro* before and after KH7 application ($n=5$). **h**, Total exploration time in NOR experiments as described in Fig. 3d. Data, mean \pm s.e.m.; * $P < 0.05$; *** $P < 0.001$; NS, not significant. For statistics, see Supplementary Tables 1-3.



Extended Data Figure 6 | Mitochondrial PKA activity mediates the effects of mtCB₁ receptors. **a**, The PKA inhibitor H89 (800 nM) blocks the effect of THC (800 nM) and WIN (100 nM) on complex-I activity in $CB_1^{+/+}$ purified brain mitochondria ($n = 3-6$). **b**, H89 blocks the effect of THC on PKA activity in $CB_1^{+/+}$ purified brain mitochondria ($n = 6$). **c**, H89 and THC have no effect on PKA activity in $CB_1^{-/-}$ brain mitochondria ($n = 3-5$). **d**, Quantification of data presented in Fig. 4c ($n = 3$), indicating a lower mitochondrial PKA content after silencing of AKAP121. **e**, **f**, Silencing of AKAP121 lowers mitochondrial PKA activity. **e**, Representative immunoblots of proteins phosphorylated by PKA (PKA-dependent phosphorylation) in 3T3 cells transfected with control vector or

shAKAP121. SHDA, succinate dehydrogenase A (a subunit of complex II). **f**, Quantification of data presented in **e** ($n = 12$). **g-i**, Silencing of AKAP121 blocks the effect of THC on cellular respiration. **g**, Silencing of AKAP121 blocks the effect of THC on basal cellular respiration in 3T3 cells ($n = 4$). **h**, Representative immunoblotting of AKAP121 expression in extracts of hippocampal primary cultures in the presence or absence of shAKAP121. **i**, shAKAP121 inhibits the effect of THC on basal cellular respiration of primary hippocampal cultures ($n = 5-10$). Data, mean \pm s.e.m.; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; NS, not significant. For statistics, see Supplementary Tables 1-3.



Extended Data Figure 7 | Activation of mtCB₁ decreases PKA-dependent phosphorylation of the complex-I subunit NDUF2, eventually mediating the effects on respiration. **a**, THC specifically decreases complex-I- but not complex-II- or complex-IV-dependent respiration in purified brain mitochondria in a CB₁-dependent manner ($n = 3-4$). **b**, Representative two-dimension (BN-PAGE/SDS-PAGE) immunoblotting (4 independent experiments) showing PKA-dependent phosphorylation of protein extracts of brain mitochondria from CB₁^{+/+} and CB₁^{-/-} mice treated with vehicle or THC (800 nM). **c**, Representative immunofluorescence of HeLa cells transfected with empty vector (mock), the mitochondrial-targeted constitutively active PKA (MLS-PKA-CA)

or the phosphomimetic mutant of NDUF2 (NDUF2-PM). Green, myc tag staining; red, mitochondrial staining (TOM20). **d**, Baseline cellular respiration of HEK293 cells transfected with plasmids expressing CB₁, CB₁ + MLS-PKA-CA or CB₁ + NDUF2-PM ($n = 4-6$). **e**, The effect of vehicle or WIN on cellular respiration in HEK293 cells as in **d** ($n = 4-6$). **f**, Representative images of hippocampi injected with AAV-GFP, AAV-MLS-PKA-CA or AAV-NDUF2-PM. Green, GFP fluorescence (left panel) or myc tag staining (central and right panel). Arrows, subunit of complex I, NDUF2. *** $P < 0.001$ between genotypes (**a**) or between vehicle and WIN (**d**); ** $P < 0.01$; *** $P < 0.001$ as compared to control. For statistics details, see Supplementary Tables 1-3.

Extended Data Table 1 | Theoretical *in silico* predictions of mitochondrial and plasma membrane localization of CB₁ and DN22-CB₁ receptor proteins, based on the programs MitoProt and PSORT

Protein	Amino acid sequence	Probability of Targeting		
		MITOCHONDRIA		PLASMA MEMBRANE
		MitoProt Prediction (24)	PSORT Prediction (25)	PSORT Prediction (25)
CB₁	MKSILDGLADTTFRITITDLLVGSNDIQYEDIKGDMSKLGYPQKFPLTSFRGSPFQEKMTAGDNSPLVPAGDTTNITEFYNKSLSSFKENEDNIQCGENFMDMECFMILNPSQQLAIAVLSLTGFTTVLENLLVLCVILHSRLRCRPSYHFIGSLAVADLLGSVIFVYSFVDFHVFHVKDSPNVFLFKLGGVTASFTASVGSFLTAIDRYISIIHRPLAYKRIVTRPKAVVAFCLMWTIAIVIAVLPLLGNCKKLQSVCSDFPLIDETYLMEFWIGVTSVLLLFIVYAYMYILWKAHSHAVRMIQRGTQKSIITSEDGKVQVTRPDQARMDIRLAKTLVLILVLLIICWGPLLAIMVYDVFGKMNKLIKTVFAFCMMLCLLNSTVNPIIYALRSKDLRHAFRSMFPSCGTAQPLDNSMGDSCLHKHANNTASMHRAAESCIKSTVKIAKVTMSVSTDTSACL	41%	47%	60%
DN22-CB₁	MGSNDIQYEDIKGDMSKLGYPQKFPLTSFRGSPFQEKMTAGDNSPLVPAGDTTNITEFYNKSLSSFKENEDNIQCGENFMDMECFMILNPSQQLAIAVLSLTGFTTVLENLLVLCVILHSRLRCRPSYHFIGSLAVADLLGSVIFVYSFVDFHVFHVKDSPNVFLFKLGGVTASFTASVGSFLTAIDRYISIIHRPLAYKRIVTRPKAVVAFCLMWTIAIVIAVLPLLGNCKKLQSVCSDFPLIDETYLMEFWIGVTSVLLLFIVYAYMYILWKAHSHAVRMIQRGTQKSIITSEDGKVQVTRPDQARMDIRLAKTLVLILVLLIICWGPLLAIMVYDVFGKMNKLIKTVFAFCMMLCLLNSTVNPIIYALRSKDLRHAFRSMFPSCGTAQPLDNSMGDSCLHKHANNTASMHRAAESCIKSTVKIAKVTMSVSTDTSACL	1%	3%	60%

Designer matrices for intestinal stem cell and organoid culture

Nikolce Gjorevski¹, Norman Sachs², Andrea Manfrin¹, Sonja Giger¹, Maïia E. Bragina¹, Paloma Ordóñez-Morán³, Hans Clevers² & Matthias P. Lutolf^{1,4}

Epithelial organoids recapitulate multiple aspects of real organs, making them promising models of organ development, function and disease^{1–3}. However, the full potential of organoids in research and therapy has remained unrealized, owing to the poorly defined animal-derived matrices in which they are grown⁴. Here we used modular synthetic hydrogel networks^{5,6} to define the key extracellular matrix (ECM) parameters that govern intestinal stem cell (ISC) expansion and organoid formation, and show that separate stages of the process require different mechanical environments and ECM components. In particular, fibronectin-based adhesion was sufficient for ISC survival and proliferation. High matrix stiffness significantly enhanced ISC expansion through a yes-associated protein 1 (YAP)-dependent mechanism. ISC differentiation and organoid formation, on the other hand, required a soft matrix and laminin-based adhesion. We used these insights to build a fully defined culture system for the expansion of mouse and human ISCs. We also produced mechanically dynamic matrices that were initially optimal for ISC expansion and subsequently permissive to differentiation and intestinal organoid formation, thus creating well-defined alternatives to animal-derived matrices for the culture of mouse and human stem-cell-derived organoids. Our approach overcomes multiple limitations of current organoid cultures and greatly expands their applicability in basic and clinical research. The principles presented here can be extended to identify designer matrices that are optimal for long-term culture of other types of stem cells and organoids.

Organoids formed by self-organizing stem cells resemble their native counterparts in cellular content, multicellular architecture and functional features. As such, they are emerging as powerful tools in basic and translational research. Intestinal organoids possess a particularly high level of multicellular organization and a wide range of potential applications⁷. When cultured under the appropriate 3D conditions, single Lgr5-expressing ISCs undergo cycles of self-renewal, differentiation and morphogenesis, and self-organize into crypt–villus domains that house both cycling ISCs and differentiated intestinal cells⁸. These organoids hold promise as models of the ISC niche, intestinal development and function⁷, as models of patient-specific epithelial diseases⁹, and as sources of tissue for autologous transplants¹⁰.

Current organoid culture methods are complicated by the nearly exclusive dependence on animal-derived hydrogels, including Matrigel¹¹ and collagen¹², as the 3D matrix. These matrices feature complex and variable compositions^{13,14}, are not conducive to controlled modifications and pose risks of immunogen and pathogen transfer, which makes them unsuitable for organoid expansion with downstream clinical applications. Synthetic ECM analogues are promising alternatives to native matrices^{5,15}. Biochemical and biophysical components can be varied, and their influence on biological processes can be tested systematically, allowing for the identification of cellular-system-specific

3D culture environments. Here we used synthetic ECMs to define the microenvironmental parameters that govern ISC behaviour and intestinal organoid formation. We used this insight to create a well-defined 3D matrix for the culture of ISCs and intestinal organoids.

To assess whether a 3D gel matrix is sufficient for ISC expansion and organoid formation, we embedded single dissociated mouse ISCs into enzymatically crosslinked polyethylene glycol (PEG) hydrogels¹⁶, which are soft and hydrated, thus mimicking the basic physical properties of Matrigel without contributing any biochemical signals. Whereas cells cultured in Matrigel divided to form lumen-containing colonies (Extended Data Fig. 1a), those embedded in soft (~300 Pa) PEG gels did not (Fig. 1a), suggesting that not only the physical support in 3D, but also biochemical signals produced by the matrix are required for ISC survival and proliferation. Indeed, enrichment of our inert PEG gels with fibronectin, laminin-111, collagen IV, hyaluronic acid and perlecan, key ECM components located at the bottom of the crypts *in vivo*^{17–21}, enhanced ISC survival and proliferation (Fig. 1a, b).

Aiming to create the most minimal, chemically defined environment, we replaced full-length fibronectin with an RGD (Arg-Gly-Asp) peptide. Like fibronectin, RGD stimulated ISC colony formation in a concentration-dependent fashion (Fig. 1c, d). Freshly isolated mouse intestinal crypts embedded in RGD-functionalized PEG gels (PEG RGD) likewise survived and formed expanding colonies (Extended Data Fig. 1b). Cells maintained in PEG RGD preserved their colony formation efficiency and Lgr5 expression after at least four passages, indicating that single ISCs can be successfully expanded within PEG RGD gels (Fig. 1e, f) without losing their organoid formation capacity, once re-embedded in Matrigel (Fig. 1g, h). Quantitative reverse transcription polymerase chain reaction (qRT-PCR) confirmed that, under ISC expansion conditions, Lgr5 was highly expressed in both PEG RGD and Matrigel (Extended Data Fig. 1c). As expected, markers of differentiated intestinal cell types became upregulated under organoid formation conditions. Notably, lysozyme appeared to be significantly expressed in Matrigel-based ISC expansion culture, indicating that colonies expanded in Matrigel contain a population of differentiated Paneth cells, which were notably absent in the PEG gels (Extended Data Fig. 1d). Hence, minimal PEG-based matrices may be more suitable for the maintenance of a pure Lgr5⁺ ISC population.

Embryonic and adult stem cells, including ISCs, are influenced by their mechanical environment^{22,23}. We increased our understanding of the mechanical regulation of ISCs by considering the effects of matrix stiffness, which have not been examined before, as performing controlled mechanical modulations *in vivo* and in Matrigel is challenging. In the presence of 1 mM RGD, matrix stiffness indeed influenced ISC colony formation (Fig. 2a, b, Extended Data Fig. 1e): ISC expansion was optimal within matrices of intermediate stiffness (1.3 kPa), whereas stem cells embedded in soft matrices (300 Pa) proliferated poorly. ISC expansion in stiff matrices required cytoskeletal tension

¹Laboratory of Stem Cell Bioengineering, Institute of Bioengineering, School of Life Sciences (SV) and School of Engineering (STI), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. ²Hubrecht Institute and University Medical Center Utrecht, Uppsalalaan 8, 3584CT Utrecht, The Netherlands. ³Swiss Institute for Experimental Cancer Research (ISREC), EPFL, Lausanne, Switzerland. ⁴Institute of Chemical Sciences and Engineering, School of Basic Science, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland.

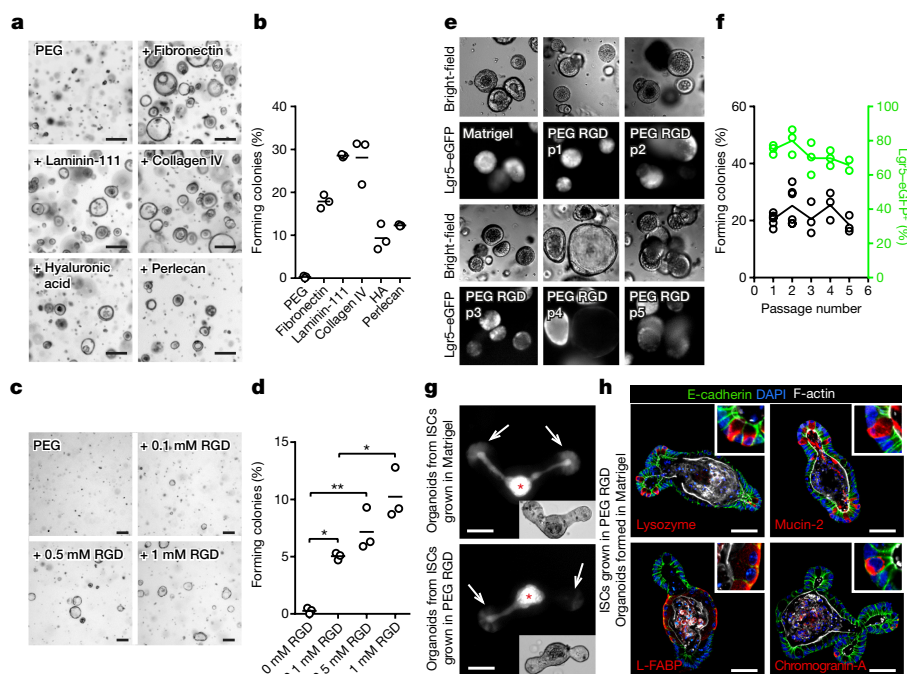


Figure 1 | Building a synthetic ISC niche. **a**, Mouse ISCs embedded in plain PEG and PEG functionalized with the indicated ECM components. **b**, Quantification of colony formation efficiency. HA, hyaluronic acid. **c**, **d**, Effect of RGD concentration on ISC colony formation in PEG (**c**) and quantification (**d**). **e**, **f**, ISC colony formation in Matrigel and PEG RGD (**e**) and quantification (**f**). **g**, Organoids formed

from ISC colonies expanded in Matrigel and PEG RGD. Asterisks indicate autofluorescence. Arrows indicate Lgr5-expressing ISCs. **h**, Organoids formed from ISC colonies expanded in PEG RGD and then transplanted into Matrigel contain differentiated intestinal cells. Graphs show individual data points derived from $n = 3$ independent experiments and means. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.01$. Scale bars, 50 μm .

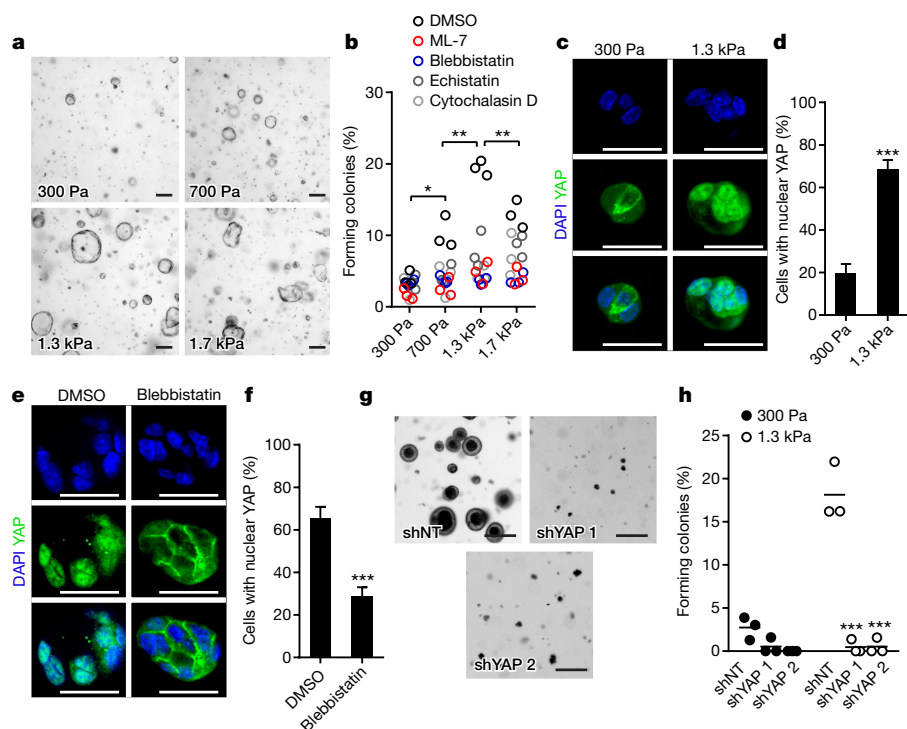


Figure 2 | Matrix mechanical properties control ISC proliferation. **a**, ISC colony formation in PEG RGD hydrogels of varying stiffness. **b**, Effect of matrix stiffness, cellular contractility and adhesion on ISC colony formation. Graphs show individual data points derived from $n = 3$ independent experiments and means. **c**, **d**, Distribution of YAP in day 1 ISC colonies (**c**) and quantification (**d**). $n = 27$ colonies (300 kPa) and $n = 21$ (1.3 kPa). Data are represented as mean \pm s.e.m. **e**, **f**, Distribution of

YAP in day 1 colonies formed in 1.3 kPa matrices treated with DMSO or blebbistatin (**e**) and quantification (**f**). $n = 27$ colonies (DMSO) and $n = 30$ (blebbistatin). Data are represented as mean \pm s.e.m. **g**, Effect of YAP knockdown on ISC colony formation (**g**) and quantification (**h**). Graphs show individual data points derived from $n = 3$ independent experiments and means. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.01$. Scale bars, 50 μm .

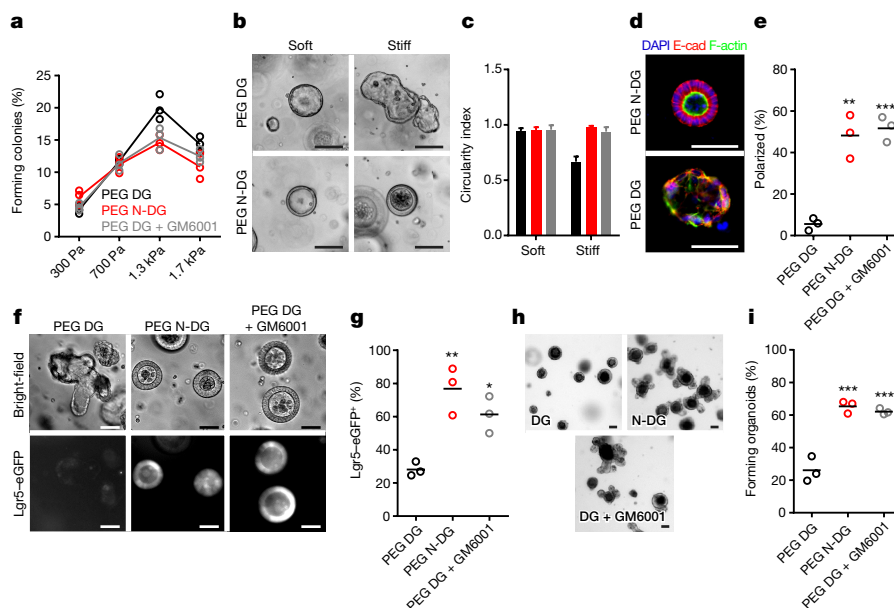


Figure 3 | ISC proteolytic activity influences proliferation, morphology and fate. **a**, Colony formation efficiency of ISCs embedded in degradable (DG) or non-degradable (N-DG) PEG gels of varying stiffness. **b**, **c**, The effect of matrix degradability on ISC colony shape (**b**) and quantification (**c**). Data are represented as mean \pm s.d. Number of colonies analysed was 13 (soft PEG DG), 25 (stiff PEG DG), 5 (soft PEG N-DG), 5 (stiff PEG N-DG), 7 (soft PEG DG + GM6001) and 8 (stiff PEG DG + GM6001). **d**, Colonies in PEG DG and PEG N-DG, stained for F-actin and

E-cadherin. **e**, Quantification of the percentage of columnar and polarized colonies. **f**, **g**, Lgr5-eGFP expression within colonies cultured in PEG DG and PEG N-DG (**f**) and quantification (**g**). **h**, **i**, Organoid formation from ISC colonies expanded in PEG DG or PEG N-DG and transferred to Matrigel (**h**) and quantification (**i**). Individual data points derived from $n = 3$ independent experiments and means are shown (**a**, **e**, **g**, **i**). * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. Scale bars, 50 μ m.

and integrin-based adhesion, as pharmacological blocking of cell contractility and integrin binding reduced colony formation (Fig. 2b).

YAP, an effector of the Hippo signalling pathway, is crucial for ISC self-renewal and expansion during intestinal tissue regeneration and organoid formation *in vitro*^{24,25}. In addition, YAP has been implicated in cellular mechanosensing and mechanotransduction, displaying enhanced nuclear translocation and activation in response to mechanical tension^{26,27}. To test whether YAP mediates the mechanical dependence of ISC expansion, we assessed its subcellular localization within colonies grown in soft matrices compared to stiff matrices. Whereas colonies in soft matrices featured predominantly cytoplasmic localization, YAP was highly nuclear within cells grown in stiff matrices (Fig. 2c, d). Pharmacological dissipation of mechanical tension led to a significant reduction in the nuclear translocation of YAP in stiff matrices (Fig. 2e, f). Likewise, controlled matrix softening (Extended Data Fig. 2a) resulted in a change in colony morphology (Extended Data Fig. 2b), cytoplasmic translocation of YAP within 4–8 h (Extended Data Fig. 2c, d), and growth arrest (Extended Data Fig. 2e). Depletion of YAP, using lentiviral delivery of short hairpin RNA (shRNA) (Extended Data Fig. 2f), led to a marked reduction in ISC colony formation within PEG hydrogels (Fig. 2g, h), as did blocking YAP activity with the small molecule verteporfin (Extended Data Fig. 2g), which suggests that YAP activity is required for ISC expansion within PEG matrices. Collectively, these data suggest that matrix stiffness controls ISC expansion by influencing YAP activity.

Noting that proteolytic remodelling is important for morphogenesis and repair²⁸, we reasoned that making the PEG RGD gels susceptible to degradation by matrix metalloproteinases would render them more physiological, and potentially enhance ISC growth. Therefore, a mutated collagen-I-derived sequence¹⁶, which can be rapidly cleaved by cell-secreted matrix metalloproteinases, was incorporated in the backbone of the polymer network. Although conferring degradability enhanced ISC colony formation in stiff matrices (Fig. 3a), the resulting colonies adopted irregular shapes (Fig. 3b, c), were depolarized and disorganized (Fig. 3d, e), and had attenuated Lgr5 expression (Fig. 3f, g) and organoid formation capacity (Fig. 3h, i).

RNA sequencing and gene set enrichment analysis revealed significant upregulation of stress and inflammation signalling-related genes within degradable matrices (Extended Data Fig. 3, Extended Data Tables 2 and 3; GEO, GSE85391). Hence, by enabling ISCs cultured within a stiff environment to engage in rapid matrix degradation, we modelled a pathological, inflammation-like state, that appears to be deleterious to stem cell maintenance.

Next, to test whether PEG RGD gels optimized for ISC expansion also support differentiation and organoid formation, we switched expanded colonies derived from Lgr5-eGFP-ires-CreERT2 mouse intestinal crypts to differentiation conditions, which led to global loss of Lgr5-eGFP and colony destruction in both degradable and non-degradable PEG RGD gels (Extended Data Fig. 4a). To shed light on the deficiencies blocking organoid formation within PEG RGD, we monitored differentiating ISC colonies in both PEG RGD and Matrigel by time-lapse imaging. In Matrigel, we observed attenuation in Lgr5 expression within 24 h (Extended Data Fig. 4b). Faint Lgr5-eGFP expression was detected at punctate locations, which formed buds by 48 h; the Lgr5-eGFP signal was then recovered in the newly formed buds. In synthetic matrices, the initial steps paralleled those in Matrigel. However, instead of forming external buds, the Lgr5-eGFP epithelial patches thickened and buckled (Extended Data Fig. 4b). Comparing the temporal profiles of YAP localization within PEG and Matrigel, we found that, within stiff matrices, the proportion of cells with nuclear YAP sharply dropped over time (Extended Data Fig. 4c). In contrast, colonies cultured in Matrigel had sustained, intermediate levels of nuclear YAP (Extended Data Fig. 4c). We proposed that, although stiff matrices initially enhance YAP activity within ISCs by promoting cellular tension^{26,27}, the resulting colony growth within a stiff confining environment may give rise to compression, which would lead to YAP inactivation, thus retarding further growth and morphogenesis^{26,29}.

We therefore sought to design mechanically dynamic matrices, which would afford the higher stiffness initially required for YAP activation and ISC expansion, but would subsequently soften to alleviate the accumulation of compressive forces, and permit *in vitro* organogenesis. Thus, we created hybrid PEG hydrogels wherein a stable polymer

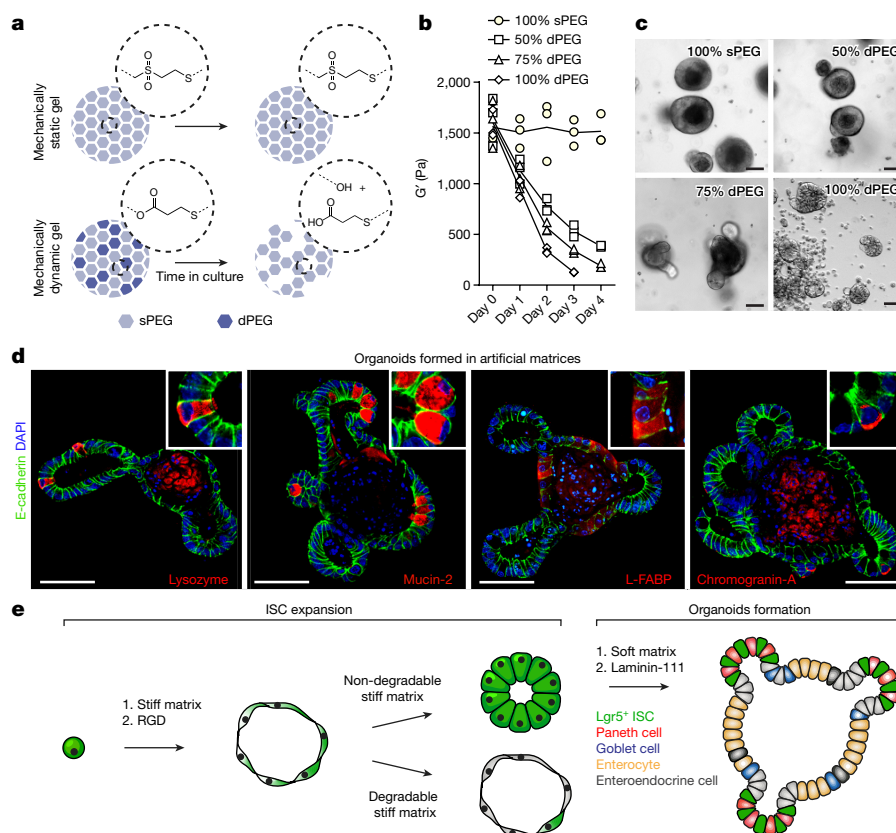


Figure 4 | Intestinal organoid formation within well-defined matrices. **a**, A schematic illustrating the chemical basis of the mechanically dynamic hydrogels. **b**, Mechanical characterization of hybrid sPEG–dPEG gels. For each condition, graph shows individual data points derived from $n = 3$ independently prepared gels and means. G' , shear modulus. **c**, Organoid

formation within matrices of varying softening rates. **d**, Organoids formed in PEG hydrogels contain differentiated intestinal cells. **e**, Summary of the rules governing different stages of organoid formation within an artificial matrix. Scale bars, 50 μm .

backbone (sPEG, mechanically static PEG) was partially replaced with hydrolytically degradable polymer (dPEG, mechanically dynamic PEG)³⁰, the presence of which afforded sustained gel softening (Fig. 4a). Varying the ratio of sPEG to dPEG within the gels allowed us to control the softening profile over time (Fig. 4b). ISC colony formation in all conditions was extensive. Upon switching the colonies to organoid formation conditions, we found that matrix softening was not sufficient to induce budding in the presence of RGD only (data not shown). In the presence of RGD and laminin-111, but not laminin-derived peptides (Extended Data Fig. 5i), the softening profile of the gels significantly influenced the emergent phenotype (Fig. 4c): organoids formed only in gels with a stiffness of around 190 Pa, suggesting that organoid formation in a minimal environment occurs in a narrow mechanical window. Immunofluorescence analysis demonstrated the presence of differentiated cells within organoids grown in minimal matrices (Fig. 4d). As predicted, at the time of differentiation, the extent of YAP activation was significantly higher in softening matrices, compared with those of constant, high stiffness (Extended Data Fig. 5j). We suggest that this difference accounts for the separate morphogenetic outcomes within the two matrices.

Finally, to test whether these matrices are also suitable for the culture of clinically relevant human counterparts, we embedded human small intestinal and colorectal cancer organoids in PEG RGD matrices. As reported before, growth was less vigorous than that of mouse cells, yet most of the embedded organoids survived and continued to expand (Extended Data Fig. 5k, l), suggesting that the PEG-based hydrogels reported here can be adapted for the culture of human organoids.

We report the first, to our knowledge, synthetic matrix for the expansion of mouse and human ISCs and a minimal matrix for the culture of mouse intestinal organoids. In addition to a final product with a

wide range of applications in research and therapy, the design process provides mechanistic insight into how the 3D microenvironment influences multiple aspects of organoid formation, including ISC proliferation, fate decisions and self-organization. YAP has been implicated in the regulation of ISC behaviour, and described as a mechanotransducer in other cellular systems—here we bridge these roles, demonstrating that the YAP-mediated control of ISC self-renewal can have a mechanical upstream component. Notably, elucidating the role of ECM components and physical parameters on ISC behaviour was made possible by the modular nature of the synthetic hydrogels.

In creating a well-defined matrix for intestinal organoids, we identified general design principles that may be adopted in building synthetic scaffolds for different organoid systems. We found that, within a minimal environment, separate stages of the organoid formation process require different microenvironments (Fig. 4e). Indeed, the native cellular environment evolves to accommodate the dynamic needs of tissues during development, homeostasis and repair. Mimicking the dynamic character of native microenvironments in designer matrices may not only replace animal-derived gels for organoid culture, but may also help to enhance the similarity of organoids to real organs.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 4 February; accepted 18 October 2016.

Published online 16 November 2016.

1. Clevers, H. Modeling development and disease with organoids. *Cell* **165**, 1586–1597 (2016).
2. Lancaster, M. A. & Knoblich, J. A. Organogenesis in a dish: modeling development and disease using organoid technologies. *Science* **345**, 1247125 (2014).

3. Sasai, Y. Cytosystems dynamics in self-organization of tissue architecture. *Nature* **493**, 318–326 (2013).
4. Fatehullah, A., Tan, S. H. & Barker, N. Organoids as an *in vitro* model of human development and disease. *Nat. Cell Biol.* **18**, 246–254 (2016).
5. Lutolf, M. P., Gilbert, P. M. & Blau, H. M. Designing materials to direct stem-cell fate. *Nature* **462**, 433–441 (2009).
6. Ranga, A. *et al.* 3D niche microarrays for systems-level analyses of cell fate. *Nat. Commun.* **5**, 4324 (2014).
7. Sato, T. & Clevers, H. Growing self-organizing mini-guts from a single intestinal stem cell: mechanism and applications. *Science* **340**, 1190–1194 (2013).
8. Sato, T. *et al.* Single Lgr5 stem cells build crypt-villus structures *in vitro* without a mesenchymal niche. *Nature* **459**, 262–265 (2009).
9. Dekkers, J. F. *et al.* A functional CFTR assay using primary cystic fibrosis intestinal organoids. *Nat. Med.* **19**, 939–945 (2013).
10. Yui, S. *et al.* Functional engraftment of colon epithelium expanded *in vitro* from a single adult Lgr5⁺ stem cell. *Nat. Med.* **18**, 618–623 (2012).
11. Kleinman, H. K. & Martin, G. R. Matrigel: basement membrane matrix with biological activity. *Semin. Cancer Biol.* **15**, 378–386 (2005).
12. Jabaji, Z. *et al.* Type I collagen as an extracellular matrix for the *in vitro* growth of human small intestinal epithelium. *PLoS One* **9**, e107814 (2014).
13. Hughes, C. S., Postovit, L. M. & Lajoie, G. A. Matrigel: a complex protein mixture required for optimal growth of cell culture. *Proteomics* **10**, 1886–1890 (2010).
14. Vukicevic, S. *et al.* Identification of multiple active growth factors in basement membrane Matrigel suggests caution in interpretation of cellular activity related to extracellular matrix components. *Exp. Cell Res.* **202**, 1–8 (1992).
15. Seliktar, D. Designing cell-compatible hydrogels for biomedical applications. *Science* **336**, 1124–1128 (2012).
16. Ehrbar, M. *et al.* Biomolecular hydrogels formed and degraded via site-specific enzymatic reactions. *Biomacromolecules* **8**, 3000–3007 (2007).
17. Benoit, Y. D., Groulx, J. F., Gagné, D. & Beaulieu, J. F. RGD-dependent epithelial cell-matrix interactions in the human intestinal crypt. *J. Signal Transduct.* **2012**, 248759 (2012).
18. Simo, P. *et al.* Changes in the expression of laminin during intestinal development. *Development* **112**, 477–487 (1991).
19. Simon-Assmann, P. *et al.* Differential expression of laminin isoforms and $\alpha 6$ - $\beta 4$ integrin subunits in the developing human and mouse intestine. *Dev. Dyn.* **201**, 71–85 (1994).
20. Wang, F. *et al.* Isolation and characterization of intestinal stem cells based on surface marker combinations and colony-formation assay. *Gastroenterology* **145**, 383–395 (2013).
21. Yamamoto, S. *et al.* Heparan sulfate on intestinal epithelial cells plays a critical role in intestinal crypt homeostasis via Wnt/ β -catenin signaling. *Am. J. Physiol. Gastrointest. Liver Physiol.* **305**, G241–G249 (2013).
22. Fernández-Sánchez, M. E. *et al.* Mechanical induction of the tumorigenic β -catenin pathway by tumour growth pressure. *Nature* **523**, 92–95 (2015).
23. Guilak, F. *et al.* Control of stem cell fate by physical interactions with the extracellular matrix. *Cell Stem Cell* **5**, 17–26 (2009).
24. Gregorieff, A., Liu, Y., Inanlou, M. R., Khomchuk, Y. & Wrana, J. L. Yap-dependent reprogramming of Lgr5⁺ stem cells drives intestinal regeneration and cancer. *Nature* **526**, 715–718 (2015).
25. Imajo, M., Ebisuya, M. & Nishida, E. Dual role of YAP and TAZ in renewal of the intestinal epithelium. *Nat. Cell Biol.* **17**, 7–19 (2015).
26. Aragona, M. *et al.* A mechanical checkpoint controls multicellular growth through YAP/TAZ regulation by actin-processing factors. *Cell* **154**, 1047–1059 (2013).
27. Dupont, S. *et al.* Role of YAP/TAZ in mechanotransduction. *Nature* **474**, 179–183 (2011).
28. Sternlicht, M. D. & Werb, Z. How matrix metalloproteinases regulate cell behavior. *Annu. Rev. Cell Dev. Biol.* **17**, 463–516 (2001).
29. Halder, G., Dupont, S. & Piccolo, S. Transduction of mechanical and cytoskeletal cues by YAP and TAZ. *Nat. Rev. Mol. Cell Biol.* **13**, 591–600 (2012).
30. Nicodemus, G. D. & Bryant, S. J. Cell encapsulation in biodegradable hydrogels for tissue engineering applications. *Tissue Eng. Part B Rev.* **14**, 149–165 (2008).

Acknowledgements We thank M. Knobloch (University of Zurich) for helpful discussions, A. Negro (EPFL) for help in the development of PEG–alginate hybrid hydrogels, D. Ossipov (Uppsala University) for providing hyaluronic acid, the Lausanne Genomic Technologies Facility (K. Harshman) for RNA-seq and D. Pioletti (EPFL) for rheometer use. N.G. was supported by an EMBO Long-Term Postdoctoral Fellowship. This work was also supported by funding from the Ecole Polytechnique Fédérale de Lausanne (EPFL). Work performed in the laboratory of H.C. was supported by the NWO Translational Adult Stem Cell Research Grant 40-41400-98-1108 and by a NWO VENI Grant 916.15.182.

Author Contributions N.G. and M.P.L. conceived the study, designed experiments, analysed data and wrote the manuscript. N.G. was involved in performing and analysing all experiments in the manuscript except for those involving human organoids. P.O.M. helped design experiments and analyse RNA-seq data. A.M. performed qPCR gene expression experiments and analysed data and produced lentiviruses. S.G. performed flow cytometry analysis of integrin expression of ISCs culture in Matrigel and PEG RGD. M.E.B. designed and characterized PEG–alg hydrogel system and helped N.G. perform experiments with ISCs in these matrices. N.S. and H.C. designed experiments and analysed data with human cells. N.S. performed experiments with human cells.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M. L. (matthias.lutolf@epfl.ch).

Reviewer Information *Nature* thanks L. Li, J. Mills and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

Data reporting. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment. No statistical methods were used to determine sample size.

Mice. Intestinal crypts were extracted from 5–10-week-old heterozygous *Lgr5*–eGFP–IRES–CreERT2 mice (Jackson Laboratory), following animal experimentation protocols prescribed by EPFL and FELASA.

Intestinal crypt isolation. Murine intestinal crypts were isolated following procedures described previously. In brief, the proximal part of the intestine was collected, opened longitudinally and washed with ice-cold PBS. The luminal side of the intestine was scraped using a glass slide to remove luminal content and villous structures. After washing with ice-cold PBS again, the intestine was cut into 2–4 mm pieces with scissors. The pieces were transferred to a 50 ml Falcon tube and further washed with cold PBS (5–10 times) with gentle vortexing. Intestinal fragments were then incubated in 20 mM EDTA/PBS for 20 min on ice. EDTA was removed, 10 ml of cold PBS was added and crypts were released by manual shaking of the suspension for 5 min. The supernatant was collected and passed through a 70- μ m strainer (BD Biosciences). The remaining tissue fragments were resuspended in 10 ml cold PBS, triturated 5–10 times and the supernatant was passed through a 70- μ m strainer. The previous step was repeated a second time. The three crypt-containing fractions were combined and centrifuged at 110g for 5 min. The pellet was resuspended in 10 ml cold Advanced DMEM/F12 (Invitrogen) and centrifuged at 84g to remove single cells and tissue debris. The resulting pellet was enriched in crypts, which were subsequently dissociated or directly embedded in PEG gels or in Matrigel (BD Biosciences; growth factor reduced, phenol red-free formulation). When needed, crypts or ISC colonies were dissociated enzymatically by incubating for 8 min at 37°C in 1 ml TrypLE Express (Life Technologies), supplemented with DNase I (2000 U ml^{−1}; Roche), 0.5 mM *N*-acetylcysteine (Sigma) and 10 μ M Y27632 (Stemgent). Undigested clusters were removed by passing the suspension through a 40 μ m strainer.

Cell culture. Freshly isolated mouse crypts or single cells from dissociated mouse ISC colonies were embedded in Matrigel or PEG gels, which were cast into 20- μ l droplets at the bottom of wells in 24-well plate. Following polymerization (15 min, 37°C), the gels were overlaid with 500 μ l of ISC expansion medium (Advanced DMEM/F12 containing Glutamax, HEPES, penicillin-streptomycin, B27, N2 (Invitrogen) and 1 μ M *N*-acetylcysteine (Sigma)), supplemented with growth factors, including EGF (50 ng ml^{−1}; R&D), Noggin (100 ng ml^{−1}; produced in-house) and R-spondin (500 ng ml^{−1}; produced in-house), and small molecules, including CHIR99021 (3 μ M; Millipore) and valproic acid (1 mM; Sigma). For single-cell culture, thiazovivin (2.5 μ M; Stemgent) was included in the medium during the first two days. To induce stem cell differentiation and organoid formation, the medium was removed, the gels were washed with PBS and fresh medium containing EGF, Noggin and R-spondin was added. Human small intestinal and colorectal cancer organoids were generated as described previously^{31,32} and grown in 20- μ l droplets of Matrigel or PEG gels overlaid with Advanced DMEM/F12 containing Glutamax, HEPES, penicillin-streptomycin, B27 (Life Technologies), Wnt3a (50% conditioned medium; produced in-house; only for small intestinal organoids), R-spondin 1 (20% conditioned medium; produced in-house), Noggin (10% conditioned medium; produced in-house), *N*-acetylcysteine (2 μ M; Sigma), Nicotinamide (10 mM; Sigma), human EGF (50 ng ml^{−1}; Peprotech), A83-01 (500 nM; Tocris), SB202190 (10 μ M; Sigma), Prostaglandin E2 (10 nM; Tocris), Gastrin (10 nM; Tocris), and Y-27632 (10 μ M; Abmole). In general, growth factors were replenished every two days, with full medium change taking place every four days. Where indicated, the following compounds were used at the specified concentrations: blebbistatin (Sigma, 12.5 μ M), ML7 (Calbiochem, 10 μ M), cytochalasin D (Merck-Millipore, 0.1 μ g ml^{−1}), echistatin (Sigma, 500 nM).

PEG, peptides and synthesis of hydrogel precursors. Vinylsulfone-functionalized 8-arm PEG (8-arm PEG-VS or sPEG) was purchased from NOF, and acrylate-functionalized 8-arm PEG (8-arm PEG-Acr or dPEG) was purchased from Creative PEGWorks. The transglutaminase (TG) factor XIII (FXIIIa) substrate peptides Ac-FKGGGPGIWGQ-ERCG-NH2 (TG-DG-Lys), Ac-FKGG-GDQGIAGF-ERCG-NH2 (TG-NDG-Lys) and H-NQEQVSPLERCGNH2 (TG-Gln) and the RGD-presenting adhesion peptide H-NQEQVSPLERGDSPG-NH2 (TG-Gln-RGD) were purchased from GL Biochem. To couple the FXIIIa substrate peptides to the 8-arm PEG-VS or 8-arm PEG-Acr, they were mixed with the PEG powder in a 1.2 stoichiometric excess (peptide-to-VS group); the combined solids were dissolved in triethanolamine (0.3 M, pH 8.0), and allowed to react for 2 h at 37°C. The reaction solution was dialysed (Snake Skin, MWCO 10K, PIERCE) against ultrapure water for 3 days at 4°C, after which the five products ((PEG-VS)-DG-Lys, (PEG-VS)-NDG-Lys, (PEG-VS)-Gln, (PEG-Acr)-NDG-Lys, (PEG-Acr)-Gln) were lyophilized. The resulting solid precursors were dissolved in ultra-pure water to make 13.33% w/v stock solutions.

Formation and dissociation of PEG hydrogels. Appropriate volumes of 13.33% w/v PEG precursor solutions were mixed in stoichiometrically balanced ratios to generate hydrogel networks of a desired final PEG content. Hydrogel formation was triggered by the addition of thrombin-activated FXIIIa (10 U ml^{−1}; Galexis) in the presence of Tris-buffered saline (TBS; 50 mM, pH 7.6) and 50 mM CaCl₂. The spare reaction volume was used for the incorporation of dissociated ISCs, fragments of human small intestinal or colorectal cancer organoids, and ECM components: TG-RGD-Gln, fibronectin (0.5 mg ml^{−1}; Invitrogen), laminin-111 (0.1 mg ml^{−1}; Invitrogen), collagen-IV (0.25 mg ml^{−1}; BD Bioscience), hyaluronic acid (0.5 mg ml^{−1}; gift from D. Ossipov, Uppsala University), perlecan (0.05 mg ml^{−1}; Sigma). Gels were allowed to crosslink by incubating at 37°C for 15 min. Dissociation and release of colonies grown in PEG for downstream cell processing or re-embedding was accomplished by enzymatic digestion of the gels. Gels were carefully detached from the bottom of the plate using the tip of a metal spatula and transferred to a 15-ml Falcon tube containing 1 ml of TrypLE Express (Life Technologies), supplemented with DNase I (2,000 U ml^{−1}; Roche), 0.5 mM *N*-acetylcysteine (Sigma) and 10 μ M Y27632 (Stemgent). Following digestion (10 min, 37°C), the cell suspension was washed with 10 ml of cold medium, passed through a 40- μ m strainer (BD Biosciences) and centrifuged at 1,200 r.p.m. for 5 min.

Softening of PEG hydrogels through ester-based hydrolysis. To form mechanically dynamic PEG hydrogels, which underwent varying extents of spontaneous softening, hybrid hydrogels were formed from both PEG-VS and PEG-Acr hydrogel precursors. Specifically, to form a fast-softening 100% Acr gel, stoichiometric quantities of (PEG-Acr)-NDG-Lys and (PEG-Acr)-Gln precursors were allowed to crosslink. A slow-softening 50% Acr gel was formed by crosslinking stoichiometric amounts of (PEG-VS)-NDG-Lys and (PEG-Acr)-Gln precursors. A 75% Acr gel with intermediate kinetics of softening was formed by crosslinking the (PEG-Acr)-Gln precursor with half of the stoichiometric equivalent of (PEG-VS)-NDG-Lys and half of the stoichiometric equivalent of (PEG-Acr)-NDG-Lys. Regardless of the relative proportions of (PEG-VS) and (PEG-Acr) precursors within the hydrogel, its overall PEG content was varied to tune its initial mechanical properties. It should be noted that, by providing an initially stiff and later a soft environment, the mechanically dynamic matrices support both ISC expansion and organoid formation. Hence, ISCs can be expanded and organoids can be formed in the same hydrogel.

PEG-alginate hydrogels for controlled matrix softening. Hybrid PEG–alginate (PEG–alg) gels were employed to induce a controlled drop in stiffness at a desired time. PEG–alg gels were formed by the simultaneous presence of activated FXIII enzyme—to drive the crosslinking of the PEG macromers—and Ca²⁺ ions, which induce the crosslinking of the alginate polysaccharides. Hybrid gels were formed by casting a solution, containing 2% (w/v) of stoichiometrically balanced (PEG-VS)-NDG-Lys and (PEG-VS)-Gln precursors, 10 mM TG-RGD-Gln, 10 U ml^{−1} FXIIIa, 0.8% (w/v) alginate (Sigma) and dissociated ISCs, within a 1% agarose/2% gelatin mould, containing 20 mM CaCl₂. The solution was incubated at 37°C for 15 min, carefully de-molded from the agarose substratum and transferred to a 12-well plate containing 1 ml of complete ISC expansion medium. Matrix softening was induced at the desired time by adding 1 U ml^{−1} alginate lyase (Sigma), and incubating for 1 h at 37°C. The digested gels were washed and transferred to freshly prepared ISC expansion medium.

Mechanical characterization of PEG hydrogels. The shear modulus of the PEG gels was determined by performing small-strain oscillatory shear measurements on a Bohlin CVO 120 rheometer. In brief, preformed hydrogel discs 1–1.4 mm in thickness were allowed to swell in complete cell culture medium for at least 3 h, and were subsequently sandwiched between the parallel plates of the rheometer. The mechanical response of the gels was recorded by performing frequency sweep (0.1–10 Hz) measurements in a constant strain (0.05) mode, at 37°C. The shear modulus (G') is reported as a measure of gel mechanical properties.

Quantification of ISC colony formation efficiency, ISC colony shape, morphology and *Lgr5* expression. To quantify the colony formation efficiency of single embedded ISCs, phase contrast z-stacks spanning the entire thickness of the cell-laden Matrigel or PEG gels were collected (Zeiss Axio Observer Z1) at 5 different locations within the gels. The Cell Counter plugin in ImageJ (NIH) was used to quantify the fraction of cells which had formed colonies at day 4 after seeding. To quantify colony circularity, phase contrast images of ISC colonies grown in the condition of interest (between 5 and 38 colonies per condition) were taken, and their contours traced manually in ImageJ. The circularity of the contours was measured using the Measure algorithm in ImageJ. To characterize the cell morphology within ISC colonies grown in different conditions, phase contrast images of at least 50 colonies were taken and the numbers of colonies containing packed columnar cells versus spread cells were counted. To quantify *Lgr5* expression within ISC colonies grown within different matrices, fluorescence images of at least 50 colonies

per condition were recorded and the number of colonies expressing Lgr5-eGFP was counted.

Laminin peptide hydrogels. To identify a short sequence that supports intestinal organoid culture, we created a library of soft ($G' = 200$ Pa) hydrogels in which binding sequences from the laminin $\alpha 1$ subunit previously shown to be biofunctional^{33,34} (Extended Data Table 4) were tethered to the PEG backbone. Embedding fragments of pre-formed organoids and screening the library for organoid survival and growth revealed that two laminin-derived peptides—AG73 and A55—significantly enhanced organoid viability and supported further growth (Extended Data Fig. 5a). Presenting these two sequences (A55 and AG73) alongside in the same gel did not appear to have an additive effect, likely owing to a redundant adhesion mechanism. Hence, we focused on the sequence with a stronger individual effect, that is, AG73 and the corresponding PEG gels (referred to as TG PEG-AG73). Varying the amount of AG73 peptide tethered to the PEG gel backbone revealed a dose-dependent effect on intestinal organoid viability and growth (Extended Data Fig. 5b, c). Despite the improved rate of survival and morphogenesis in TG PEG-AG73 matrices compared with plain PEG or PEG RGD, the process was significantly less efficient compared with Matrigel, and morphological differences were apparent. Keeping in mind that the effect of AG73 was concentration-dependent, we attributed these differences to a potentially sub-optimal AG73 ligand density within the synthetic system. By design, there is an upper limit to the concentration of tethered factors that can be incorporated into the PEG system used thus far in the study; exceeding this limit disrupts the structural integrity of the gels. To overcome this limitation and enhance the biofunctionality of the synthetic matrix by increasing the concentration of AG73 ligands, we turned to chemically crosslinked PEG gels. Here, vinyl sulfone (VS)-conjugated 4-arm PEG precursors are covalently linked into solid hydrogels through Michael-type addition between VS groups and the thiols of a short crosslinker containing two cysteine residues. To incorporate the AG73 ligand at a high density, we designed a crosslinker in which the AG73 sequence was flanked by two short cysteine-containing sequences. The resulting gels (hereafter referred to as MT PEG-AG73) presented the AG73 ligand at a concentration of 3.1 mM, thus significantly surpassing the highest concentration achieved in the enzymatically crosslinked matrices. Embedding intestinal organoid fragments into MT PEG-AG73 revealed that the percentage of tissues that remained viable and continued to undergo morphogenesis approached that observed in Matrigel (Extended Data Fig. 5e). To verify the maintenance of ISCs within the organoids grown in MT PEG-AG73, we embedded tissues extracted from the Lgr5-eGFP reporter mouse and monitored eGFP expression. We observed that Lgr5-eGFP was expressed in the expected pattern: localized to the crypt-like buds of the organoids (Extended Data Fig. 5f). The fraction of organoids expressing the marker was significantly higher than in those cultured in TG PEG-AG73, and at least as high as in organoids cultured in Matrigel (Extended Data Fig. 5g). We also confirmed that the organoids cultured in MT PEG-AG73 were polarized and contained differentiated cells (Extended Data Fig. 5h).

Immunofluorescence analysis. ISC colonies or organoids embedded in Matrigel or PEG gels were fixed with 4% paraformaldehyde (PFA) in PBS (30 min, room temperature). The fixation process typically led to complete degradation of the Matrigel. Suspended tissues were collected and centrifuged (800 r.p.m., 5 min) to remove the PFA, washed with ultra-pure water and pelleted. Following resuspension in water, the organoids were spread on glass slides and allowed to attach by drying. Attached organoids were rehydrated with PBS. Following fixation, organoids embedded in PEG or spread on glass were permeabilized with 0.2% Triton X-100 in PBS (1 h, room temperature) and blocked (10% goat serum in PBS containing 0.01% Triton X-100) for at least 3 h. Samples were subsequently incubated overnight at 4°C with phalloidin-Alexa 546 (Invitrogen) and primary antibodies against lysozyme (1:50; Thermo Scientific PA1-29680), mucin-2 (1:50; Santa Cruz sc-15334), chromogranin-A (1:50; Santa Cruz sc-13090), L-FABP (1:50; Santa Cruz sc-50380) and YAP1 (1:50; Santa Cruz sc-101199) diluted in blocking buffer. After washing with PBS for at least 3 h, samples were incubated overnight at 4°C with secondary antibody Alexa 647 goat- α -rabbit (1:1000 in blocking solution; Invitrogen). Following extensive washing, stained organoids were imaged in epifluorescence (Zeiss Axio Observer Z1) or confocal (Zeiss LSM 710) mode. Alternatively, ISC colonies or organoids cultured in PEG were released from the hydrogel before PFA fixation, by incubating the gels with 1 mg ml⁻¹ Dispase (Gibco) for 7 min at 37°C. The released colonies or organoids were fixed with PFA in suspension, and attached to glass coverslips, as described above.

Immunohistochemical analysis of human organoids. Human organoids were fixed in 10% neutral buffered formalin, washed with PBS, dehydrated, and embedded in paraffin. Sections were stained with H&E or Ki67 antibody (1:250; Monosan).

Lentivirus production and transduction. Lentiviral particles encoding for shRNA recognizing YAP (two sequences validated for knockdown, purchased from Sigma) or the pLKO.1-puro Non-Target control shRNA (Sigma) were generated in HEK 293T cells, using third generation lentivirus packaging vectors. Transfection was carried out using the X-tremeGENE HP Transfection kit (Roche). After 48 h, the supernatant was collected, filtered and ultracentrifuged at 50,000g for 2 h at 20°C. The resulting pellet was resuspended in PBS and stored at -80°C. Lentiviral infection of ISCs was performed by dissociating the ISC colonies (described above), resuspending the resulting single cells in ice-cold liquid Matrigel, containing 10 μ M Y27632 and the concentrated lentiviral particles at a dilution of 1:10. ISCs were incubated with viral particles in a liquid suspension for 45 min on ice. The suspension was subsequently cast into droplets and allowed to form gels, which were overlaid with ISC expansion medium. The embedded ISCs proceeded to form colonies, which carried the transgene encoded by the virus. The cells were allowed to recover and form colonies for 36 h, after which they were dissociated and encapsulated within PEG hydrogels or used for quantification of knockdown efficiency by qPCR.

Quantitative real-time PCR (qPCR). ISC colonies or organoids grown in Matrigel or PEG gels were dissociated as described above, and RNA was extracted using an RNeasy Micro Kit (Qiagen). cDNA was synthesized using the iScript cDNA Synthesis Kit (Bio-Rad). qPCR was carried out using the Power SYBR Green PCR Master Mix (Applied Biosystems) and the primers listed Extended Data Table 1.

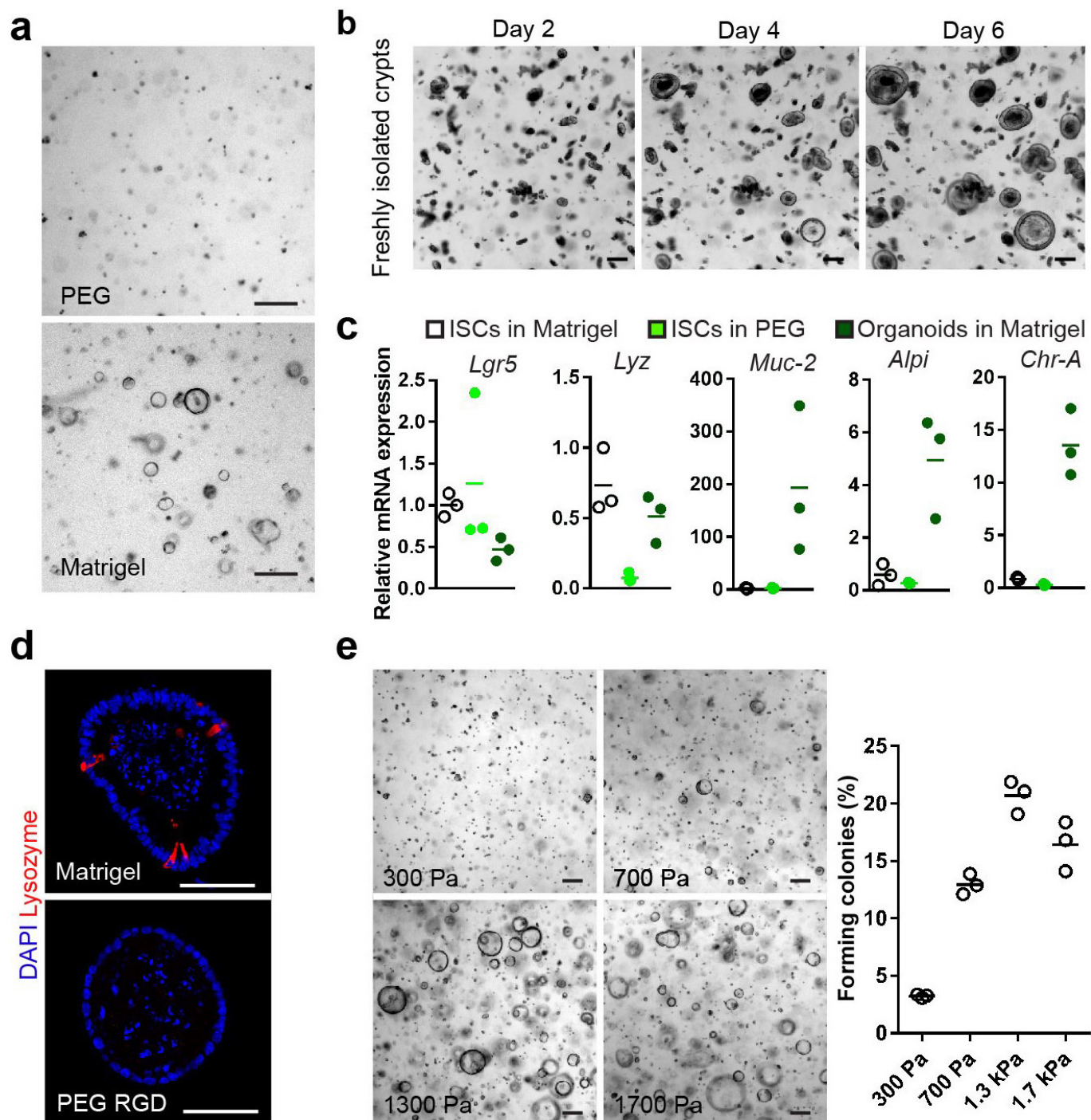
RNA-seq: library preparation, sequencing, data processing, statistical analysis, interpretation and accession numbers. RNA-seq libraries were prepared using 500 ng of total RNA and the Illumina TruSeq Stranded mRNA reagents (Illumina; San Diego, California, USA) on a Sciclone liquid handling robot (PerkinElmer; Waltham, Massachusetts, USA) using a PerkinElmer-developed automated script. Cluster generation was performed with the resulting libraries using the Illumina TruSeq SR Cluster Kit v4 reagents and sequenced on the Illumina HiSeq 2500 using TruSeq SBS Kit v4 reagents. Sequencing data were processed using the Illumina Pipeline Software version 1.84. Purity-filtered reads were adapters and quality trimmed with Cutadapt and filtered for low complexity with seq_crums (v. 0.1.8). Reads were aligned against the *Mus musculus*.GRCm38.82 genome using STAR³⁵. The number of read counts per gene locus was summarized with htseq-count³⁶ using the *Mus musculus*.GRCm38.82 gene annotation. Quality of the RNA-seq data alignment was assessed using RSeQC³⁷. Reads were also aligned to the *Mus musculus*.GRCm38.82 transcriptome using STAR and the estimation of the isoforms abundance was computed using RSEM³⁸. Statistical analysis was performed for genes in the R software package. Genes with low counts were filtered out according to the rule of 1 count per million in at least 1 sample. Library sizes were scaled using TMM normalization in the EdgeR package³⁹ and log-transformed with the limma voom function⁴⁰. Differential expression was computed with limma⁴¹ by fitting paired samples data into a linear model and performing all pairwise comparisons. To control for false discovery and multiple testing, we computed an adjusted *P* value, using the Benjamini-Hochberg method. Gene set expression analysis was performed with the freely available GSEA software⁴² (Broad Institute), using differential expression values and pre-defined gene signatures as inputs. In particular, to test for upregulation of stress-related genes, the MSigDB gene set 'BIOCARTA_STRESS_PATHWAY' was used. We checked for upregulation of colon cancer-related genes by using the inflammatory colon cancer signature, as identified by Sadanandam *et al.*⁴³. Functional annotation and gene ontology analysis of significantly enriched gene sets was conducted using the MetaCore software. The accession numbers for the gene expression profiles described here are GEO: GSE85391.

Statistical analysis and sample information. Statistically significant differences between the means of two groups were assessed by using a Student's *t*-test, whereas data containing more than two experimental groups were analysed with a one-way ANOVA followed by a Bonferroni's multiple comparison test. All statistical analyses were performed in the GraphPad Prism 6.0 software.

Data availability. RNA sequencing data that support the findings of this paper have been deposited to the Gene Expression Omnibus (GEO) public repository (GSE85391; <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE85391>). Source Data for Figs. 1–4 and Extended Data Figs 1–5 are provided with the paper. All additional relevant data are available upon request from the corresponding author.

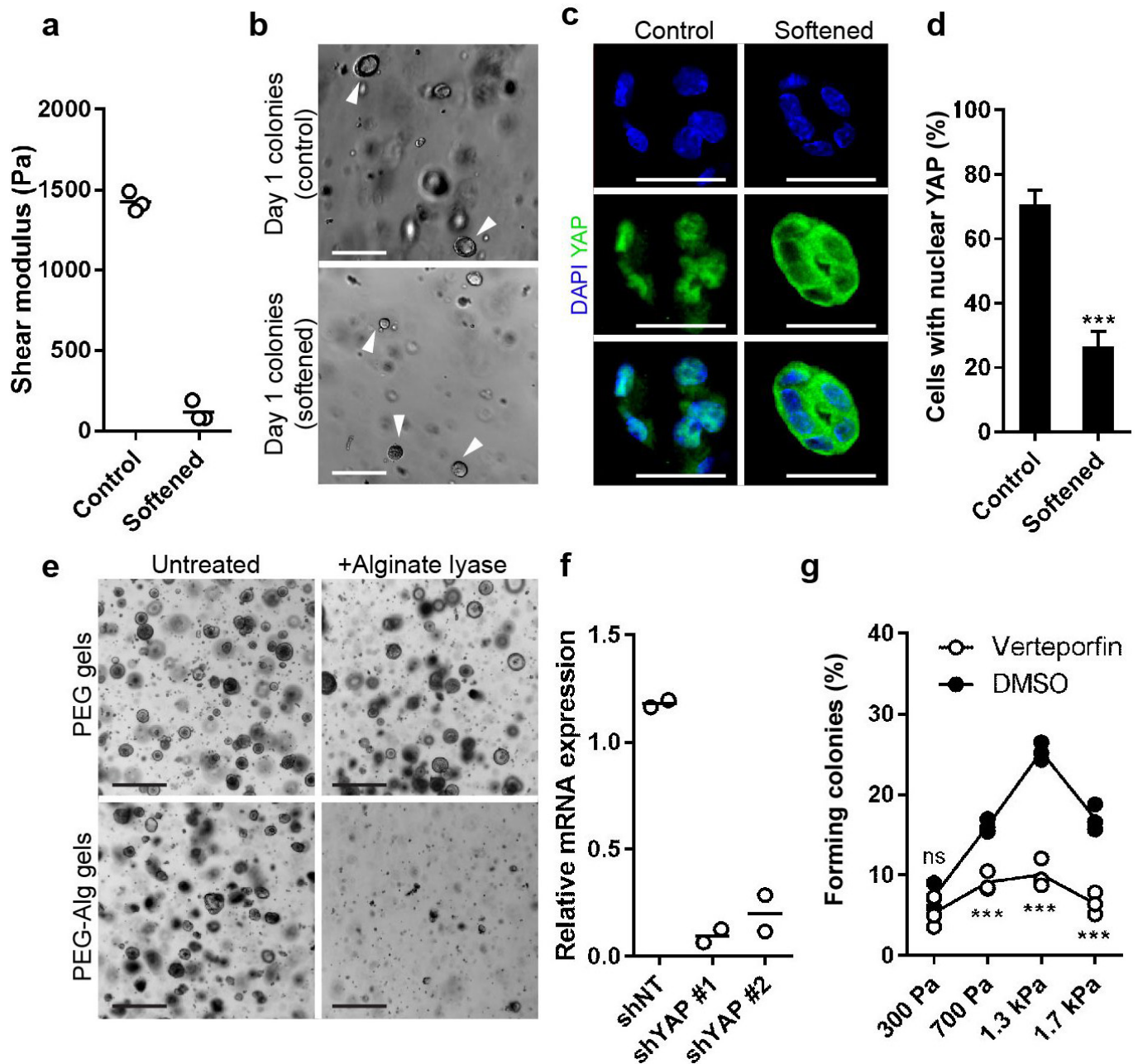
- Matano, M. *et al.* Modeling colorectal cancer using CRISPR-Cas9-mediated engineering of human intestinal organoids. *Nat. Med.* **21**, 256–262 (2015).
- Sato, T. *et al.* Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. *Gastroenterology* **141**, 1762–1772 (2011).

33. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
34. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
35. Wang, L., Wang, S. & Li, W. RSeQC: quality control of RNA-seq experiments. *Bioinformatics* **28**, 2184–2185 (2012).
36. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
37. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
38. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).
39. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
40. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
41. Sadanandam, A. *et al.* A colorectal cancer classification system that associates cellular phenotype and responses to therapy. *Nat. Med.* **19**, 619–625 (2013).
42. Nomizu, M. *et al.* Identification of cell binding sites in the laminin α 1 chain carboxyl-terminal globular domain by systematic screening of synthetic peptides. *J. Biol. Chem.* **270**, 20583–20590 (1995).
43. Nomizu, M. *et al.* Cell binding sequences in mouse laminin α 1 chain. *J. Biol. Chem.* **273**, 32491–32499 (1998).



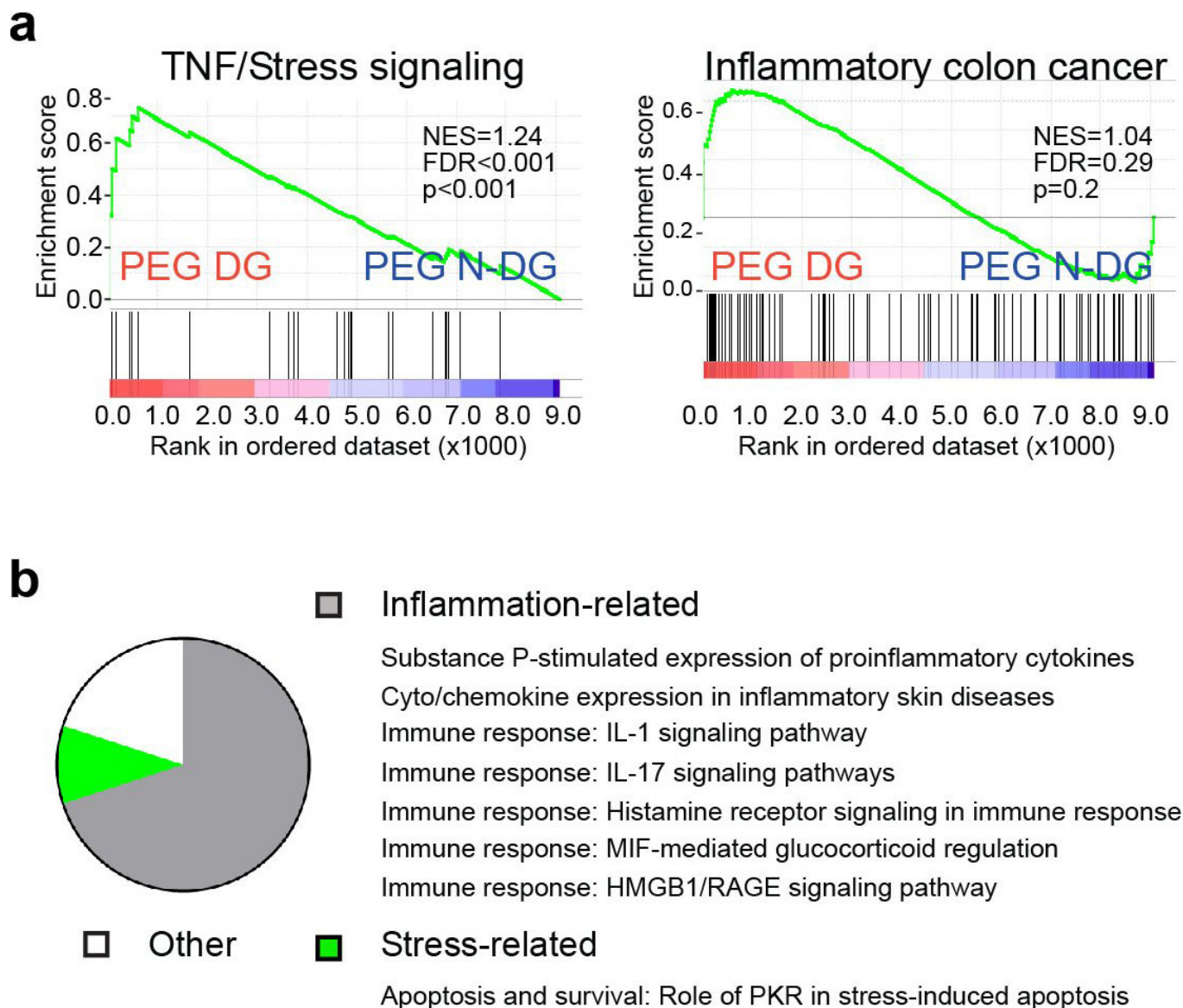
Extended Data Figure 1 | Primary crypt culture in synthetic matrices, characterization of ISCs grown in PEG RGD culture with or without CHIR99021 and valproic acid. **a**, ISCs cultured in Matrigel and unmodified PEG gels for 24 h. **b**, ISC colonies formed from freshly isolated mouse intestinal crypts embedded in PEG RGD. **c**, Relative mRNA levels of intestinal genes, quantified by qPCR. **d**, ISC colonies cultured under

self-renewal conditions in Matrigel, but not PEG gels, contain lysozyme-expressing Paneth cells. **e**, Stiffness-dependent colony formation and quantification of ISCs cultured with EGF, noggin, R-spondin and Wnt3a. Graphs show individual data points derived from $n = 3$ independent experiments and means. Scale bars, 50 μm .



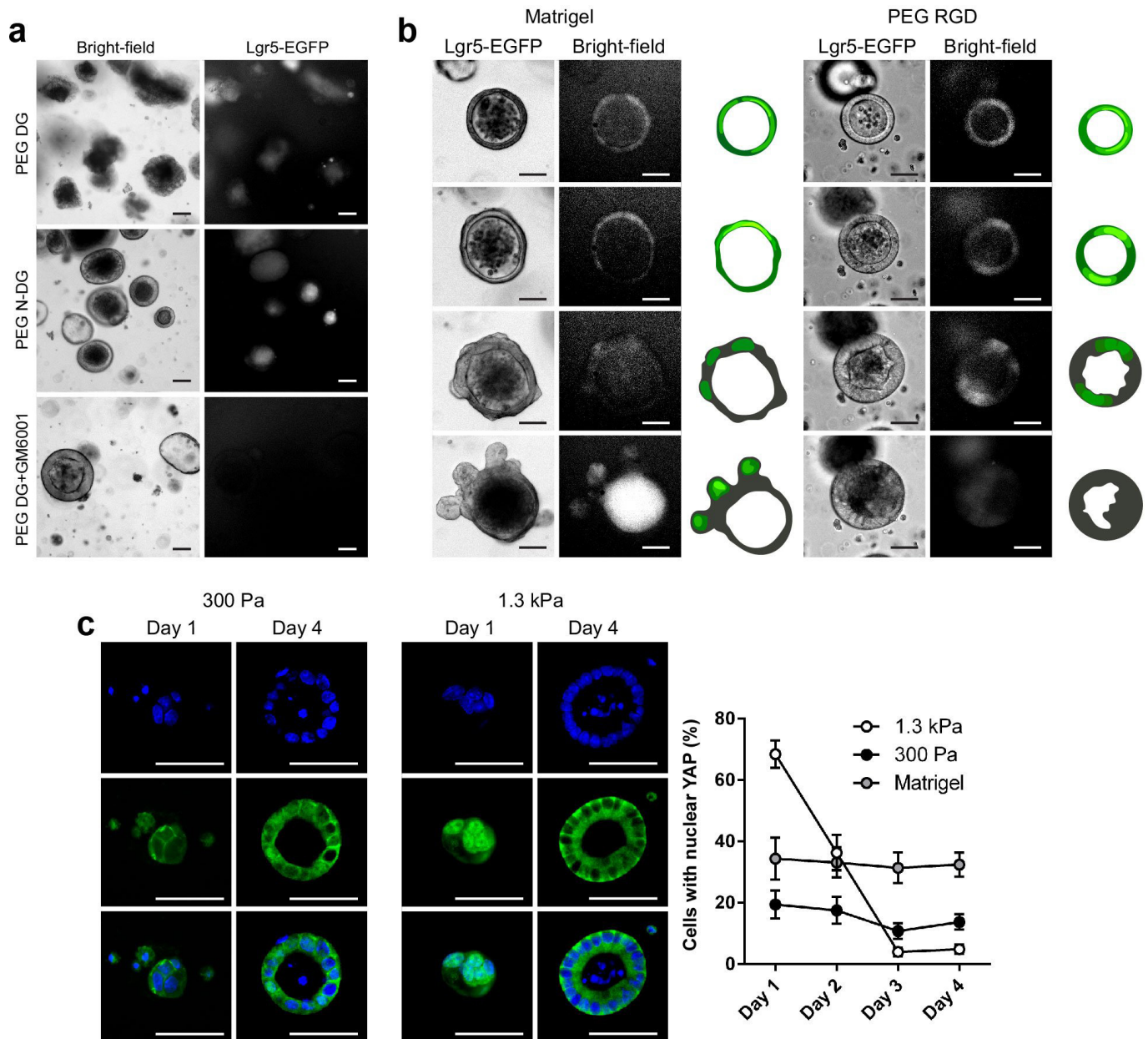
Extended Data Figure 2 | Effect of controlled matrix softening on ISC colony morphology, YAP activity and growth. **a**, Mechanical properties of control and softened PEG alginate (PEG–alg) hybrid gels obtained by selective degradation of the alginate network. Graph shows individual data points derived from 3 independently prepared gels. **b**, Morphology of day 1 colonies in control and softened PEG–alg gels. **c**, **d**, Distribution of YAP in day 1 colonies in control and softened PEG–alg gels (**c**) and quantification (**d**). $n = 28$ colonies (control) and $n = 31$ (softened). Data

are represented as mean \pm s.e.m. **e**, Alginate lyase treatment does not affect colony growth in PEG RGD gels. Softening of PEG–alg gels by alginate-lyase-mediated digestion blocks colony growth. **f**, Quantification of shRNA-mediated knockdown of YAP. Graph shows individual data points derived from 2 independent experiments and means. **g**, Effect of verteporfin on ISC colony formation in PEG RGD. Graph shows individual data points derived from 3 independent experiments and means. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. Scale bars, 50 μ m.



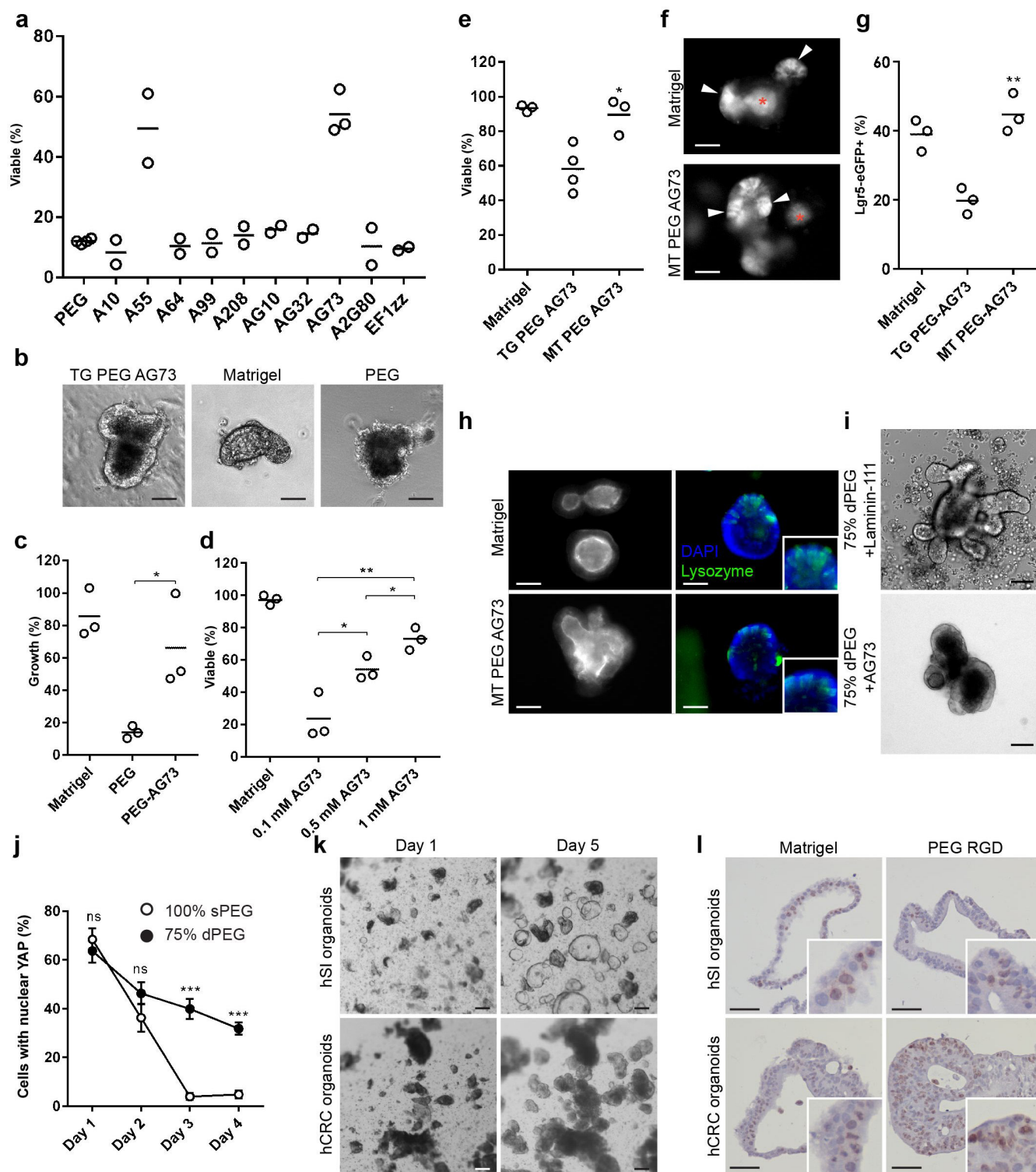
Extended Data Figure 3 | Stiff degradable matrices induce an inflammation-like state in ISCs. **a**, Gene set enrichment analysis (GSEA) comparing RNA-seq gene expression data of ISCs cultured in degradable compared to non-degradable matrices to published gene signatures

(see Methods for details). **b**, Functional annotation of signalling pathway significantly upregulated in degradable matrices (details and statistics shown in Extended Data Tables 2, 3).



Extended Data Figure 4 | Comparison of organoid formation and YAP activity in Matrigel and PEG RGD. **a**, Organoid formation does not occur within synthetic PEG RGD matrices. ISC colonies cultured in non-degradable (PEG N-DG) and degradable (PEG DG) PEG RGD matrices, with or without GM6001 for 4 d, and subsequently cultured under organoid formation conditions for 2 d. **b**, Time-course analysis of morphology and Lgr5-eGFP expression during organoid formation. ISC colonies formed in Matrigel and PEG hydrogels were monitored for 48 h, following a switch to differentiation and organoid formation conditions.

c, Localization of YAP at day 1 and day 4 in ISC colonies formed in 300 Pa and 1.3 kPa PEG RGD gels. **d**, Quantification of nuclear translocation of YAP as a function of time in 300 Pa, 1.3 kPa PEG RGD and Matrigel. Data are shown as means \pm s.e.m. To assess the extent of YAP nuclear translocation, $n = 21$ (day 1, 1.3 kPa); 27 (day 2, 1.3 kPa); 22 (day 3, 1.3 kPa); 22 (day 4, 1.3 kPa); 6 (day 2, 300 Pa); 23 (day 3, 300 Pa); 22 (day 4, 300 Pa); 30 (day 1, Matrigel); 28 (day 2, Matrigel); 30 (day 3, Matrigel); 30 (day 4, Matrigel). * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. Scale bars, 50 μ m.



Extended Data Figure 5 | Effect of laminin-derived peptides on organoid formation, and human organoid expansion in PEG RGD.

a, Effect of different laminin-111-derived sequences on intestinal organoid viability. **b**, Morphology of intestinal organoids grown in Matrigel, plain PEG or PEG-AG73 gels. **c**, AG73-conjugated PEG matrices significantly enhance the growth of intestinal organoids. **d**, Effect of AG73 on intestinal organoid viability and growth is concentration-dependent. **e**, Quantification of intestinal organoid viability in Matrigel, TG PEG-AG73 and MT PEG-AG73. **f**, Morphology and Lgr5-eGFP expression in intestinal organoids grown in Matrigel and MT PEG-AG73 gels. **g**, Quantification of Lgr5-eGFP expression in intestinal organoids expanded in Matrigel, TG PEG-AG73 and MT PEG-AG73. **h**, Establishment of apicobasal polarity and presence of Paneth (lysozyme) cells within organoids grown in MT PEG-AG73. **i**, AG73 peptides are

not capable of supporting differentiation and organoid formation from ISC colonies, whereas full-length laminin-111 is. **j**, Quantification of nuclear translocation of YAP as a function of time in cells cultured within mechanically stable and softening gels. Data are shown as means \pm s.e.m. To assess the extent of YAP nuclear translocation, $n = 21$ (day 1, 100% sPEG), 27 (day 2, 100% sPEG), 22 (day 3, 100% sPEG), 22 (day 4, 100% sPEG), 30 (day 1, 75% dPEG), 22 (day 3, 100% sPEG), 22 (day 4, 100% sPEG), 30 (day 1, 75% dPEG), 30 (day 2, 75% dPEG), 30 (day 3, 75% dPEG), and 30 (day 4, 75% dPEG) colonies were analysed. **k**, **l**, Phase contrast images (**k**) and Ki67 expression (**l**) of human ISC colonies and human patient-derived colorectal cancer organoids grown in PEG RGD. Graphs show individual data points derived from $n = 2$ (**a**), $n = 3$ (**c–e**, **g**) independent experiments and means. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. Scale bars, 50 μ m.

Extended Data Table 1 | Primers used for qPCR analysis of intestinal gene expression

Mouse gene	Forward primer	Reverse primer
<i>Lgr5</i>	ATTCGGTGCATTTAGCTTGG	CGAACACCTGCGTGAATATG
<i>Alpi</i>	GCCTATCTCTGTGGGGTCAA	TTTCTTGGCACGGTACATCA
<i>Chga</i>	AAGGTGATGAAGTGCGTCCT	GGTGTCGCAGGATAGAGAGG
<i>Muc2</i>	CCGACTTCAACCCAAGTGAT	GAGCAAGGGACTCTGGTCTG
<i>Lyz1</i>	GTCACACTTCCTCGCTTTCC	TGGCTTTGCTGACTGACAAG
<i>Gapdh</i>	ATCCTGCACCACCAACTGCT	GGGCCATCCACAGTCTTCTG
<i>Yap1</i>	TCCTCCTTTGAGATCCCTGA	GCCATGTTGTTGTCTGATCG

Extended Data Table 2 | Functional annotation of genes significantly upregulated in degradable PEG RGD matrices versus non-degradable PEG RGD matrices, in terms of cellular pathways

Pathway	Count	%	p value	FDR
Immune response_MIF-mediated glucocorticoid regulation	4	17.4	1.7E-05	4.1E-03
Protein folding and maturation_Posttranslational processing of neuroendocrine peptides	5	17.4	2.3E-05	4.1E-03
Immune response_Substance P-stimulated expression of proinflammatory cytokines via MAPKs	4	10.0	2.1E-04	1.9E-02
Immune response_IL-1 signaling pathway	4	9.3	2.3E-04	1.9E-02
Immune response_Histamine H1 receptor signaling in immune response	4	9.1	3.2E-04	1.9E-02
Muscle contraction_Relaxin signaling pathway	4	8.3	3.2E-04	1.9E-02
PDE4 regulation of cyto/chemokine expression in inflammatory skin diseases	4	8.3	3.8E-04	1.9E-02
Apoptosis and survival_Role of PKR in stress-induced apoptosis	4	8.0	4.8E-04	1.9E-02
Immune response_HMGB1/RAGE signaling pathway	4	7.5	4.8E-04	1.9E-02
Immune response_IL-17 signaling pathways	4	7.5	7.7E-04	2.7E-02

Performed using the Metacore software package.

Extended Data Table 3 | Functional annotation of genes significantly upregulated in degradable PEG RGD matrices versus non-degradable PEG RGD matrices, in terms of cellular processes

Process	Count	%	p value	FDR
Chemotaxis	7	5.1	6.1E-05	6.2E-03
Inflammation_Histamine signaling	8	3.8	1.5E-04	6.2E-03
Inflammation_Neutrophil activation	8	3.7	1.6E-04	6.2E-03
Inflammation_Innate inflammatory response	7	3.9	3.3E-04	7.6E-03
Development_Neurogenesis_Synaptogenesis	7	3.9	3.3E-04	7.6E-03
Muscle contraction	6	3.5	1.7E-03	2.8E-02
Cell adhesion_Platelet-endothelium-leucocyte interactions	6	3.4	1.7E-03	2.8E-02
Transport_Potassium transport	6	3.1	3.0E-03	3.8E-02
Muscle contraction_Relaxin signaling	4	4.9	3.1E-03	3.8E-02
Immune response_Innate immune response to RNA viral infection	4	4.8	3.4E-03	3.8E-02

Performed using the Metacore software package.

Extended Data Table 4 | Laminin-derived peptides

Abbreviation	Sequence
AG73	Ac-NQEQVSPLRKRLQVQLSIRT-NH ₂
AG10	Ac-NQEQVSP LNRWHSYITRFG-NH ₂
AG32	Ac-NQEQVSP LTWYKIAFQRNRK-NH ₂
A10	Ac-NQEQVSP LGTNNWWQSPSIQN-NH ₂
A13	Ac-NQEQVSP LRQVFQVAYIIKA-NH ₂
A55	Ac-NQEQVSP LGGFLKYTVSYDI-NH ₂
A64	Ac-NQEQVSP LRDQLMTVLANVT-NH ₂
A2G80	Ac-NQEQVSP LVQLRNGFPYFSY-NH ₂
A203	Ac-NQEQVSP LMEQANLLLDRL-NH ₂
A208	Ac-NQEQVSP LAASIKVAVSADR-NH ₂
A99	Ac-NQEQVSP LAGTFALRGDNPQG-NH ₂
EF ^{tz}	Ac-NQEQVSP LATLQLQEGRLHF(Nle)FDLGKGR-NH ₂
bis-Cys AG73	Ac-GCRERKRLQVQLSIRTERCG-NH ₂

See refs 33, 34.

Different tissue phagocytes sample apoptotic cells to direct distinct homeostasis programs

Ryan J. Cummings^{1,2}, Gaetan Barbet^{1,2}, Gerold Bongers^{1,2}, Boris M. Hartmann³, Kyle Gettler⁴, Luciana Muniz^{1,2,5}, Glaucia C. Furtado^{1,2}, Judy Cho^{1,2,6}, Sergio A. Lira^{1,2,7} & J. Magarian Blander^{1,2,5,7,8†}

Recognition and removal of apoptotic cells by professional phagocytes, including dendritic cells and macrophages, preserves immune self-tolerance and prevents chronic inflammation and autoimmune pathologies^{1,2}. The diverse array of phagocytes that reside within different tissues, combined with the necessarily prompt nature of apoptotic cell clearance, makes it difficult to study this process *in situ*. The full spectrum of functions executed by tissue-resident phagocytes in response to homeostatic apoptosis, therefore, remains unclear. Here we show that mouse apoptotic intestinal epithelial cells (IECs), which undergo continuous renewal to maintain optimal barrier and absorptive functions³, are not merely extruded to maintain homeostatic cell numbers⁴, but are also sampled by a single subset of dendritic cells and two macrophage subsets within a well-characterized network of phagocytes in the small intestinal lamina propria^{5,6}. Characterization of the transcriptome within each subset before and after *in situ* sampling of apoptotic IECs revealed gene expression signatures unique to each phagocyte, including macrophage-specific lipid metabolism and amino acid catabolism, and a dendritic-cell-specific program of regulatory CD4⁺ T-cell activation. A common 'suppression of inflammation' signature was noted, although the specific genes and pathways involved varied amongst dendritic cells and macrophages, reflecting specialized functions. Apoptotic IECs were trafficked to mesenteric lymph nodes exclusively by the dendritic cell subset and served as critical determinants for the induction of tolerogenic regulatory CD4⁺ T-cell differentiation. Several of the genes that were differentially expressed by phagocytes bearing apoptotic IECs overlapped with susceptibility genes for inflammatory bowel disease⁷. Collectively, these findings provide new insights into the consequences of apoptotic cell sampling, advance our understanding of how homeostasis is maintained within the mucosa and set the

stage for development of novel therapeutics to alleviate chronic inflammatory diseases such as inflammatory bowel disease.

Clearance of apoptotic epithelial cells within the respiratory, colonic and post-weaning mammary epithelium can be conducted by aptly positioned neighbouring epithelial cells, which serve as non-professional phagocytes^{1,3}. To examine whether apoptotic IECs are also recognized by professional phagocytes within the small intestinal lamina propria (SILP), we generated mice that express transgenic diphtheria toxin receptor (DTR) fused to enhanced green fluorescent protein (eGFP), driven by the epithelium-specific villin promoter (VDTR mice). This enabled the experimental induction of apoptosis and allowed for tracking of apoptotic cell phagocytosis by acquisition of eGFP. The villin promoter drove transgene expression in IECs of the small and large intestine (Extended Data Fig. 1a, b). We observed no gross histological changes within the small or large intestine of VDTR mice relative to C57BL/6J (B6) controls (Extended Data Fig. 1c, d). Extensive eGFP expression co-localized with epithelial pan-cytokeratin and the actin cytoskeletal binding agent phalloidin throughout the small and large intestinal epithelia (Fig. 1a and Extended Data Fig. 1e–i). Injection of VDTR mice with 10 ng g⁻¹ diphtheria toxin induced IEC death throughout the villi; dying IECs in control mice injected with phosphate buffered saline (PBS) were noted only at villi tips, characteristic of natural IEC turnover³ (Extended Data Fig. 1j).

Lowering the dose of diphtheria toxin to 2 ng g⁻¹ showed no evidence of causing epithelial erosion, villus atrophy or inflammatory cell infiltration over time (Extended Data Figs 1c, right versus middle panels, 2a, b). Expression of inflammatory *Il6*, *Il1b* and *Tnf* genes was not induced in the ileum 4 h after administration of either 2 or 10 ng g⁻¹ of diphtheria toxin. However, upregulation of these pro-inflammatory genes was observed 16 h after administration of 10 ng g⁻¹ diphtheria toxin (Fig. 1b). We observed no bacterial translocation to

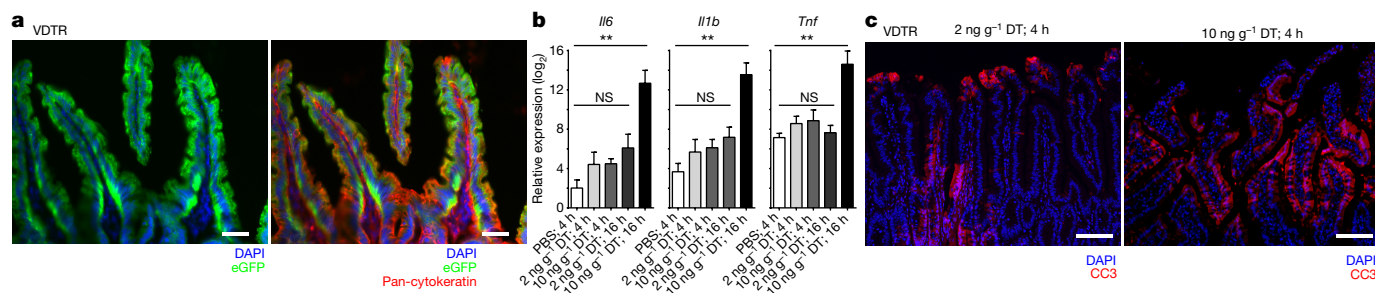


Figure 1 | A novel mouse model for inducing apoptosis of IECs under non-inflammatory conditions. **a**, Immunofluorescence for indicated markers on small intestine cryo-sections. **b**, qRT-PCR on VDTR ileum represents at least four independent experiments in duplicate. $n = 4$ mice

per group. One-way ANOVA; $**P < 0.01$, $*P < 0.05$. NS, not significant. Data are mean \pm s.e.m. **c**, Immunofluorescence for cleaved caspase 3 (CC3) on small intestine paraffin sections 4 h after administration of 2 or 10 ng g⁻¹ diphtheria toxin (DT). Scale bars, 50 (**a**) and 100 (**c**) μ m.

¹Immunology Institute, Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA. ²Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA. ³Department of Neurology, Center for Translational Systems Biology, Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA. ⁴Department of Genetics, Yale School of Medicine, New Haven, Connecticut 06520, USA. ⁵Graduate School of Biological Sciences, Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA. ⁶Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA. ⁷Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA. ⁸Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA. [†]Present address: The Jill Roberts Institute of Inflammatory Bowel Disease, Division of Gastroenterology, Joan and Sanford I. Weill Department of Medicine, Department of Microbiology and Immunology, Weill Cornell Medicine, Cornell University, New York, New York 10021, USA.

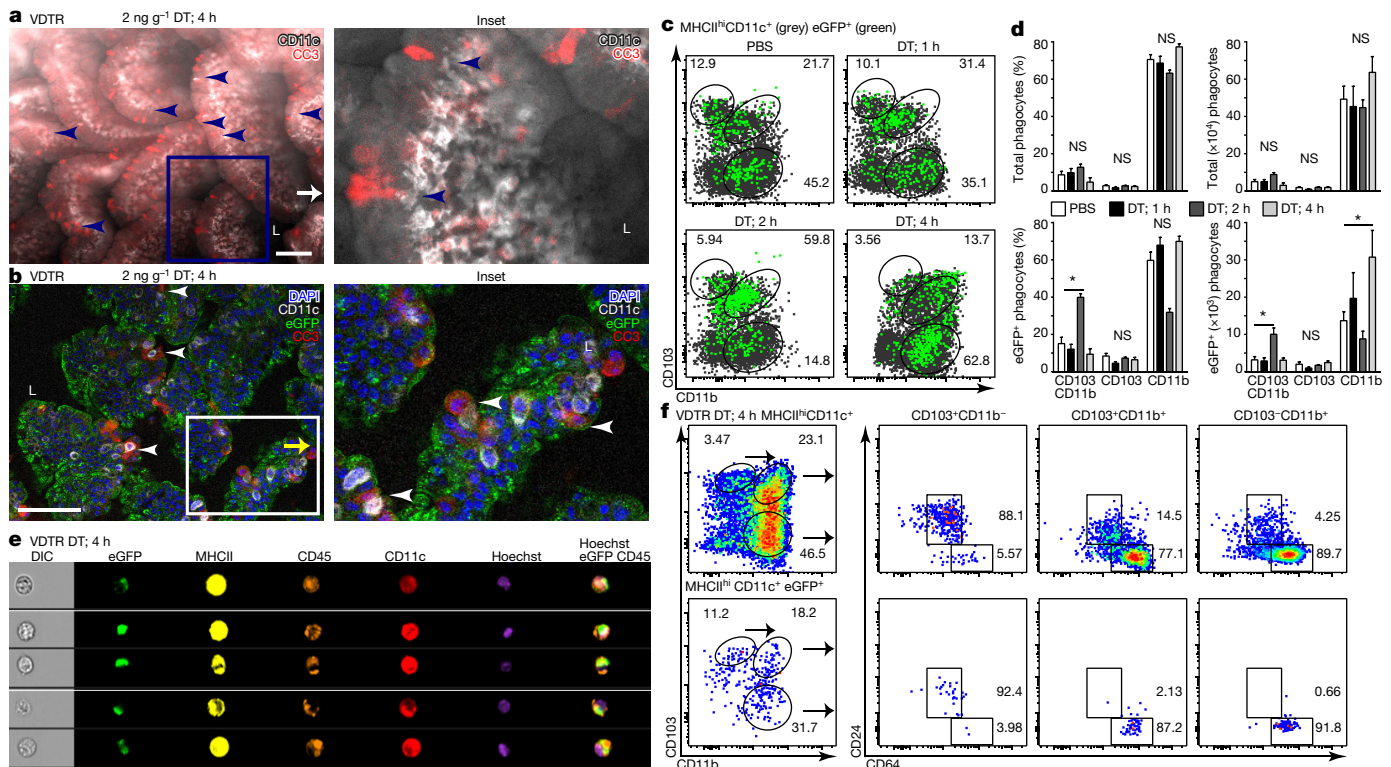


Figure 2 | A single CD103⁺ dendritic cell subset and two macrophage subsets sample apoptotic intestinal epithelial cells. **a**, **b**, Conventional (without eGFP overlaid) (**a**) and confocal (**b**) whole-mount microscopy of VDTR ileum. Scale bars, 100 μm. Arrowheads in **a** indicate phagocytes near CC3⁺ IECs. L, lumen. **c**, Flow cytometry analysis of total (grey) and eGFP⁺ (green) phagocytes pre-gated on live CD45⁺MHCII^{hi}CD11c⁺ cells. **d**, Percentage and absolute number of phagocytes from **c**. **e**, ImageStream

the intestinal lamina propria after treatment with either dose of diphtheria toxin for 4 h, in contrast to 10 ng g⁻¹ diphtheria toxin at 24 h or with 3% dextran sodium sulphate (DSS) (Extended Data Fig. 2c, d). Staining for cleaved caspase-3 (CC3), a marker of early apoptosis, was significantly increased in a dose-dependent manner within the terminal ileum of diphtheria-toxin-treated compared to PBS-treated VDTR mice (Fig. 1c and Extended Data Fig. 2e, f). We thus chose 2 ng g⁻¹ as the diphtheria toxin dose concentration that would increase the likelihood of observing phagocytic sampling of apoptotic IECs without eliciting inflammation or epithelial barrier disruption.

Using whole-mount microscopy on excised small intestine tissue, we localized CC3 labelling to eGFP⁺ IECs and CD11c expression to phagocytes, which appeared centrally within villi and proximally to the CD31⁺ vasculature (Extended Data Fig. 3a, b). We detected numerous CC3⁺ IECs in the small intestine of diphtheria-toxin-treated VDTR mice (Fig. 2a and Extended Data Fig. 3c). CD11c⁺ cells extended dendrites towards the base of apoptotic IECs upon diphtheria toxin treatment, into the space of the eGFP signal and not the lumen (as happens during soluble antigen retrieval⁶) (Fig. 2a and Extended Data Fig. 3d). Some dendrites ended with globular shapes, characteristic of phagocyte sampling⁶ (Extended Data Fig. 3e), and were observed in the same optical confocal section as CC3⁺ IECs (Fig. 2b and Extended Data Fig. 3f). CC3⁺ IECs were also increased in diphtheria-toxin-treated mice (Fig. 2b) relative to control (Extended Data Fig. 3g, h), consistent with the quantification in Extended Data Fig. 2f.

We monitored apoptotic IEC sampling by tracking the appearance of eGFP within phagocytes. We set sub-gates by flow cytometry within haematopoietic SILP cells (Extended Data Fig. 4a) from VDTR mice, compared to B6 mice (Extended Data Fig. 4b). All intestinal CD45⁺MHCII⁺CD11c⁺ phagocyte subsets, divided on the basis of surface expression of the integrins αE (CD103) and αM (CD11b)⁵,

of SILP VDTR cells. Data represent three independent experiments. DIC, differential interference contrast microscopy. **f**, Flow cytometry analysis of total phagocytes with corresponding eGFP⁺ cells. Data are mean ± s.e.m., $n = 3$ B6 mice; $n = 6$ VDTR mice treated with 2 ng g⁻¹ diphtheria toxin for 1–4 h; $n = 9$ VDTR mice treated with PBS; one-way ANOVA; * $P < 0.05$; NS, not significant. Data in **c**, **e**, and **f**, represent SILP cells from VDTR mice, at least three independent experiments.

contained the eGFP label in VDTR mice, regardless of diphtheria toxin treatment. This indicated constitutive apoptotic IEC sampling, which was not observed for monocytes and granulocytes (Extended Data Fig. 4b–e). The number of eGFP⁺ phagocytes was significantly increased 4 h after diphtheria toxin administration when compared to control (Extended Data Fig. 4f). Of all eGFP⁺ phagocytes from diphtheria-toxin-treated mice (Extended Data Fig. 4c), CD103⁺CD11b⁺ and CD103⁻CD11b⁺ (CD11b⁺) phagocytes were significantly increased at 2 h and 4 h, respectively, relative to PBS-treated mice (Fig. 2c, d). The total number of phagocytes was not increased relative to PBS-treated mice, suggesting that the increase in eGFP⁺ phagocytes was not due to an influx of cells, but rather an increase in phagocytes carrying apoptotic IEC cargo (Fig. 2c, d). ImageStream analyses showed eGFP⁺ inclusions within MHCII^{hi}CD45⁺CD11c⁺Hoechst⁺ phagocytes (Fig. 2e), indicating internalization of apoptotic bodies or degradation of formerly engulfed apoptotic IECs.

Assessment of the expression of CD64, a marker of macrophages in the mouse intestine⁶, and CD24, expression of which defines dendritic cells⁶, revealed that CD103⁺ phagocytes were predominantly CD24⁺CD64⁻, defined as dendritic cells⁶ (Fig. 2f). Minor populations of CD24⁺CD64⁻ dendritic cells were also present within the CD103⁺CD11b⁺ and CD11b⁺ phagocytes, which were otherwise predominantly comprised of CD24⁻CD64⁺ macrophages (Fig. 2f and Extended Data Fig. 4g). The CD64⁺CD103⁺CD11b⁺ macrophage population was most notable as it has not previously been reported⁸, although its CD103 expression may be transient⁵. Of CD103⁺CD11b⁺ and CD103⁻CD11b⁺ phagocytes, only CD24⁻CD64⁺ macrophages contained the eGFP label (eGFP⁺), whereas CD24⁺CD64⁻ dendritic cells did not (eGFP⁻) (Fig. 2f and Extended Data Fig. 4g). Notably, only the CD103⁺ population of CD24⁺CD64⁻ dendritic cells were eGFP⁺ (Fig. 2f and Extended Data Fig. 4g). Thus, of the five phenotypically

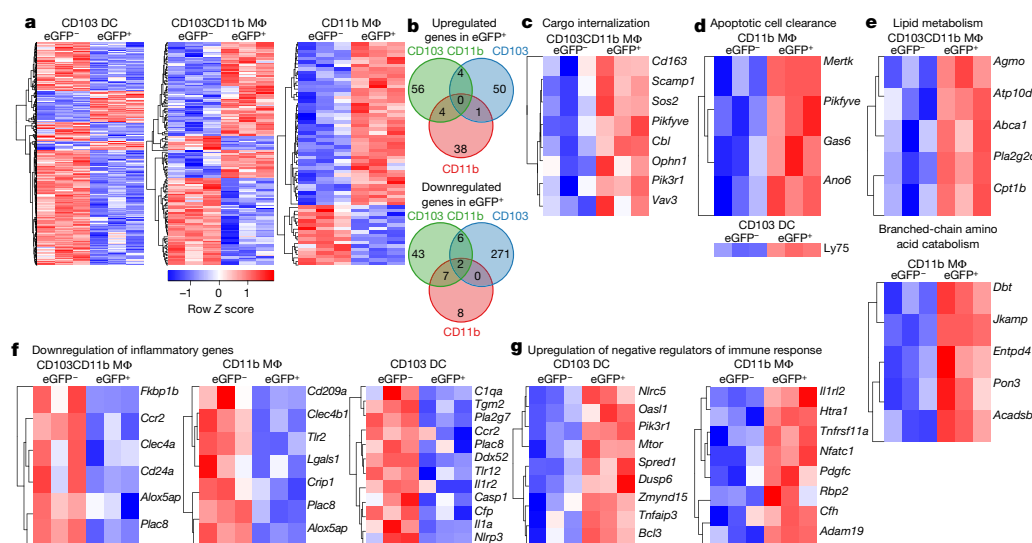


Figure 3 | Sampling apoptotic IECs *in situ* triggers a transcriptional 'suppression of inflammation' signature distinct for macrophages (M Φ) and CD103⁺ dendritic cells (DC). **a, c–g.** Hierarchical clustering of differentially expressed genes upregulated by at least 1.2 fold (ANOVA ($q < 0.05$) and Tukey's honest significant difference (HSD) post-hoc test ($P < 0.05$); $-1.2 > \text{fold-change} > 1.2$); expressed by row Z scale and organized according to Gene Ontology as indicated. **b.** Venn diagrams of genes in **a**. Data represent five independent experiments, four mice per experiment, three biological replicates.

distinct SILP phagocytes, only CD11b⁺CD103⁺ (CD103⁺) dendritic cells and two macrophage subsets, CD103⁺CD11b⁺ and CD11b⁺, sample apoptotic IECs (Fig. 2f and Extended Data Fig. 4h).

We next determined the transcriptional profiles that result from apoptotic IEC sampling by isolating the eGFP⁺ and equivalent eGFP[−] populations of each subset (Extended Data Fig. 5a) 4 h after administration of 2 ng g^{−1} diphtheria toxin, conditions that maintain a non-inflammatory state (Fig. 1 and Extended Data Fig. 2). Principal component analysis of gene expression revealed distinct clustering into dendritic cells and macrophages, and further clustering based on eGFP content (Extended Data Fig. 5b–f). Comparing gene expression profiles of eGFP⁺ and eGFP[−] phagocytes revealed predominant transcriptional downregulation in CD103⁺ dendritic cells and upregulation in macrophages after apoptotic IEC sampling (Fig. 3a). CD103⁺ dendritic cells had the most differentially expressed genes (334) followed by CD103⁺CD11b⁺ (122) and CD11b⁺ (60) macrophages (Fig. 3a, b). Each phagocyte upregulated unique sets of genes in response to apoptotic IECs, and only two genes, *Plac8* and *Itgb7* (Supplementary Information shows all gene names) were downregulated in all three of the groups, suggesting that apoptotic IEC sampling generates unique transcriptomes within each phagocytic cell subset.

Several unique pathways were induced in apoptotic IEC-bearing eGFP⁺ macrophages, including those that facilitate internalization of cargo and clearance of apoptotic cells (Fig. 3c, d). CD103⁺CD11b⁺ and CD11b⁺ macrophages, which sampled the bulk of apoptotic IECs (Fig. 2c, d), significantly upregulated a distinct array of genes involved in apoptotic cell clearance, including *Gas6*, the scavenger receptor *Cd163*, and TAM family receptor *Mertk* (encoding MER)¹, whereas CD103⁺ dendritic cells upregulated the endocytic receptor *Ly75* (encoding CD205), which detects apoptotic cell ligands⁹ (Fig. 3d). Higher surface expression of MER in macrophages and CD205 in dendritic cells correlated with higher eGFP content, irrespective of diphtheria toxin treatment, validating the use of low doses of diphtheria toxin to increase apoptotic IEC numbers without major steady-state perturbations (Extended Data Fig. 6a). Other genes whose expression correlated with eGFP content were differentially modulated in macrophages and CD103⁺ dendritic cells, including *Cx3cr1*, TAM family receptors *Axl* and *Tyro3*, and Stabilin family receptors *Adgrb1* (also known as *Bail*) *Stab1* and *Timd4* (Extended Data Fig. 6b–e, confirming similar profiles at 6 h after diphtheria toxin administration), whereas *Cd300a* remained unchanged (Extended Data Fig. 6f). Both macrophage subsets expressed *Pikfyve*, which is essential for phagosome maturation¹⁰ (Fig. 3c, d). Increased expression of genes involved in lipid processing and branched-chain amino acid catabolism was notable in apoptotic IEC-bearing CD103⁺CD11b⁺ and CD11b⁺ macrophages, respectively (Fig. 3e), and may reflect the increased catabolism of internalized apoptotic cargo metabolites. CD103⁺CD11b⁺ macrophages upregulated

Abca1 expression, which increases cholesterol efflux and has been associated with engulfment of apoptotic cells and liver X receptor α function, which upregulates *Mertk* expression¹. Similarly, CD11b⁺ macrophages upregulated *Acadslb* (Fig. 3e and Extended Data Fig. 6d).

Profiling genes that were downregulated in response to apoptotic IEC sampling revealed two notable transcriptional signatures: the downregulation of inflammatory genes and suppression of the immune response (Fig. 3f and Extended Data Fig. 6g–i). Different sets of these genes were significantly downregulated in CD103⁺CD11b⁺ and CD11b⁺eGFP⁺ macrophages, including genes that encode the pattern recognition receptors *Clec4a*, *Clec4b1*, *Cd209a* and *Tlr2* (Fig. 3f). *Alox5ap*, which facilitates inflammatory leukotriene biosynthesis, was significantly downregulated in both CD103⁺CD11b⁺ and CD11b⁺eGFP⁺ macrophages (Fig. 3f and Extended Data Fig. 6d, h) reflecting previous findings that show suppression of 5-lipoxygenase by apoptotic cells¹¹. CD103⁺ dendritic cells also significantly downregulated several innate immune genes including *Tlr12*, inflammasome-related *Nlrp3*, *Casp1* and *Il1a*¹² (Fig. 3f), as well as MAPK/ERK signalling (*Lrrk2*, *Map3k4*, *Fos*) (Extended Data Fig. 6g, i) and *Nlrp3*, *Fos* at 6 h after diphtheria toxin administration (Extended Data Fig. 6e). Classical and alternative complement genes *C1qa* and *Cfp*, respectively, were also downregulated in CD103⁺ dendritic cells (Fig. 3f), which may further dampen NLRP3 inflammasome activation¹³. Additionally, expression of *Cfh*, encoding a regulator of complement activation, was upregulated in eGFP⁺CD11b⁺ macrophages (Fig. 3g and Extended Data Fig. 6i). Also notable was the significant upregulation of negative regulators of the immune response (Fig. 3g). CD103⁺ dendritic cells upregulated: *Nlrp5* and *Tnfrsf11a* (A20), which negatively regulate the NLRP3 inflammasome, type-I interferon (IFN) signalling and NF- κ B activity¹⁴; *Oas1*, which inhibits IRF7, a type-I IFN transcription factor¹⁵; and *Spred1*, which inhibits MAPK signalling¹⁶ (Extended Data Fig. 6i, e, at 6 h after diphtheria toxin administration). Coupled with unchanged or downregulated gene expression profiles for inflammatory cytokines and co-stimulatory receptors in eGFP⁺ relative to eGFP[−] counterparts after diphtheria toxin treatment (Extended Data Fig. 7), the overall phagocytic response to apoptotic IEC sampling *in situ* results in genome-wide transcriptional suppression of a set of inflammatory genes that is unique to each CD11c⁺ subset.

In contrast to macrophages, eGFP⁺CD103⁺ SILP dendritic cells showed a distinct transcriptional profile dominated by a 'regulatory CD4⁺ T-cell activation' signature (Fig. 4a, Extended Data Fig. 6e at 6 h post-diphtheria toxin). These dendritic cells significantly upregulated genes that are required for regulatory CD4⁺ T-cell (T_{reg}) recruitment, including chemokines *Ccl22* and *Ccl17* (ref. 17), and genes that promote T_{reg}-mediated suppression—including the leucine-rich membrane protein *Lrrc32*, which retains latent TGF β on T_{reg} cells¹⁸

Table 1 | Apoptotic IEC sampling induces transcriptional modulation of IBD susceptibility genes in intestinal lamina propria phagocytes

CD103 ⁺ CD11b ⁺ MΦ				CD103 ⁺ dendritic cell			
Gene	Regulation	Gene	Regulation	Gene	Regulation	Gene	Regulation
<i>Il12b</i>	up	<i>Il12b</i>	up	<i>Sept1</i>	down	<i>Hhex</i>	down
<i>Trpm2</i>	up	<i>Lrrc32</i>	up	<i>Rassf5</i>	down	<i>Cast</i>	down
<i>Tmem132a</i>	up	<i>Pfkfb3</i>	up	<i>Ptpn1</i>	down	<i>Nusap1</i>	down
<i>Lsp1</i>	down	<i>Gca</i>	up	<i>Kif11</i>	down	<i>Il1r2</i>	down
<i>Rpl3</i>	down	<i>Cd40</i>	up	<i>Gpr183</i>	down	CD11b⁺ MΦ	
<i>Mrpl20</i>	down	<i>Tnfrsf3</i>	up	<i>Slc25a28</i>	down	Gene	Regulation
<i>Ak3</i>	down	<i>Spred1</i>	up	<i>Fuca2</i>	down	<i>Il1r2</i>	up
<i>Ctsv</i>	down	<i>Lrrk2</i>	down	<i>Mospd3</i>	down	<i>Lsp1</i>	down
<i>Sept1</i>	down	<i>Plau</i>	down	<i>Ikzf1</i>	down	Differentially expressed	
<i>Dclre1b</i>	down	<i>Mrpl20</i>	down	<i>Psmc3</i>	down	Phagocyte	Total
<i>S100a10</i>	down	<i>Fos</i>	down	<i>Acyp1</i>	down	CD103⁺ CD11b⁺ MΦ	Up
<i>Tnni2</i>	down	<i>Notch2</i>	down	<i>Sult1a1</i>	down	CD103⁺ DC	Down
<i>S100a11</i>	down	<i>Donson</i>	down	<i>Bcap29</i>	down	CD11b⁺ MΦ	

Genes exhibiting upregulation or downregulation in each SILP phagocyte population upon apoptotic IEC sampling were overlapped with IBD susceptibility loci, which were downloaded from the GEO (accession number GSE57945, <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE57945>)²⁵. Genes in bold are shared. Total numbers of differentially expressed IBD genes is shown in the bottom right corner. MΦ, macrophage.

(Fig. 4a). *Aldh1a2*, which catalyses the synthesis of retinoic acid to further generate T_{reg} in the periphery¹⁹, was also significantly upregulated (Fig. 4a, Extended Data Fig. 8a), as were genes required for T_{reg} expansion including the suppressive co-stimulatory molecule *Cd274* (encoding PD-L1)³ (Fig. 4a), whose surface expression on eGFP⁺CD103⁺ dendritic cells after diphtheria toxin administration was the same as that of untreated VDTR mice (Extended Data Fig. 8b). *Relb* and *Cd40*, involved in inducing T_{reg} differentiation²⁰, were also significantly upregulated (Fig. 4a). Thus, apoptotic IEC sampling by SILP CD103⁺ dendritic cells sets the stage for instructing differentiation of T_{reg} cells, which are important mediators of intestinal tolerance and homeostasis¹⁹.

Highest expression of the chemotactic receptor *Ccr7* in the SILP was observed in eGFP⁺CD103⁺ dendritic cells (Extended Data Fig. 8c, 6e at 6h), suggesting that CD103⁺ dendritic cells function as migratory phagocytes that traffic apoptotic IEC cargo from the lamina propria to the mesenteric lymph nodes (MLN). Migratory cells within the MLN of VDTR and B6 mice were similar, but eGFP⁺ phagocytes, which comprised approximately 10% of all migratory cells, were noted only in MLN of VDTR mice and without the need for treatment with diphtheria toxin, indicating the shuttling of apoptotic IEC cargo to the MLN under steady-state conditions (Fig. 4b and Extended Data Fig. 8d, e, B6 mice). Of the migratory subsets, only CD103⁺CD11b⁺CD24⁺ dendritic cells were eGFP⁺, comprising approximately 25% of all CD103⁺ migratory cells (Fig. 4b and Extended Data Fig. 8e), whereas MLN-resident CD103⁺ dendritic cells were eGFP⁺ (Extended Data Fig. 8f–k).

Within the MLN, migratory eGFP⁺CD103⁺ dendritic cells had a transcriptional profile (Fig. 4c) similar to that of SILP eGFP⁺CD103⁺ dendritic cells (Fig. 4a). Several transcripts affecting T_{reg} cells were additionally increased in eGFP⁺ relative to eGFP⁺CD103⁺ MLN dendritic cells, including *Ido1* and *Il10* (ref. 20). Whereas levels of *Tgfb1* transcripts, which encode a protein that facilitates T_{reg} expansion¹⁹, were similar (Extended Data Fig. 8l, m), those of *Itgb8*, which encodes a protein that activates TGFβ, were increased by approximately 10- and 20-fold in eGFP⁺ relative to eGFP⁺CD103⁺ MLN and SILP dendritic cells, respectively (Fig. 4c and Extended Data Fig. 6e). This equated to an approximately 4-fold increase from the SILP to the MLN. T-cell co-stimulatory receptors (CD274, CD40) (Fig. 4a, c and Extended Data Fig. 8b, n), and secreted signals (retinoic acid (via *Aldh1a2*), CCL22), were also increased relative to eGFP⁺CD103⁺ MLN dendritic cells (Fig. 4a, c). Consistent with this profile, eGFP⁺CD103⁺ dendritic cells induced an approximately 4-fold increase in T_{reg} cells *ex vivo*, as marked by the T_{reg}-specific transcription factor FOXP3, compared to eGFP⁺ counterparts (26.5% versus 6.6%, respectively)

and without added TGFβ (Fig. 4d). No T_{reg} differentiation occurred when TGFβ was added alone without dendritic cells (Fig. 4d). Unlike FOXP3⁺ T cells, these induced FOXP3⁺ T_{reg} cells expressed the high-affinity IL-2 receptor CD25 (ref. 19; Fig. 4e). The number of T_{reg} cells also increased *in vivo* following a single diphtheria toxin injection, as shown by increased CD45⁺CD3⁺CD4⁺FOXP3⁺ peripherally induced T_{reg} cells marked by lack of Helios expression or transcription factors specific for other T-helper cell subsets (Fig. 4f, g and Extended Data Fig. 9). Despite a predominant specificity of small intestine T_{reg} cells for dietary antigens²¹, specificities for self-antigens remain possible. Thus, although CD103⁺ dendritic cells migrate to MLN regardless of the content of apoptotic IECs (Fig. 4b), only CD103⁺ dendritic cells carrying apoptotic IECs excelled at *de novo* T_{reg}-cell differentiation.

The transposition of differentially modulated genes with the 163 susceptibility loci identified through genome-wide association studies on inflammatory bowel disease (IBD)²² showed modulation of 41 IBD susceptibility genes upon apoptotic IEC sampling, with the majority expressed in CD103⁺ dendritic cells and some select genes shared between macrophages and the CD103⁺ dendritic cells and CD103⁺CD11b⁺ macrophages (Table 1). Although no single gene was expressed in all three populations, CD103⁺ dendritic cells and CD103⁺CD11b⁺ macrophages shared downregulation of *Mrpl20* and *Sept1* and upregulation of *Il12b*, encoding the p40 subunit shared by IL-12 and IL-23—and in which single nucleotide polymorphisms are associated with Crohn's disease in Korean populations²³. Macrophage subsets shared the downregulation of *Lsp1*, whereas CD103⁺ dendritic cells upregulated *Pfkfb3*. *Lsp1* is involved in activation of PPARγ that enhances apoptotic cell phagocytosis and, together with *Pfkfb3*, suppresses diet-induced intestinal inflammation²⁴. How allelic polymorphisms within IBD susceptibility loci affect genes involved in intestinal tolerance and homeostasis is not known. However, modelling the colonic epithelial ulcerations and cellular infiltration characteristic of IBD through the treatment of VDTR mice with DSS (Extended Data Fig. 10a–c) showed downregulation of many of the T_{reg}-cell activation genes differentially expressed both upon apoptotic IEC sampling by migratory eGFP⁺CD103⁺ MLN dendritic cells, such as *Aldh1a2*, *Atrn*, *Ccl22*, *Cd274*, *Ido1*, *Il10* and *Relb*, and relative to those from untreated mice (Extended Data Fig. 10d). CD103⁺ dendritic cells were the only phagocytic subset in DSS-treated VDTR mice that delivered apoptotic IEC cargo to the MLN (Extended Data Fig. 10e), but many T_{reg} activation genes were no longer upregulated in these cells within the MLN, with expression levels similar or even reduced relative to eGFP⁺ counterparts (Extended Data Fig. 10f). Taken together, these data suggest that inflammation alters the outcome

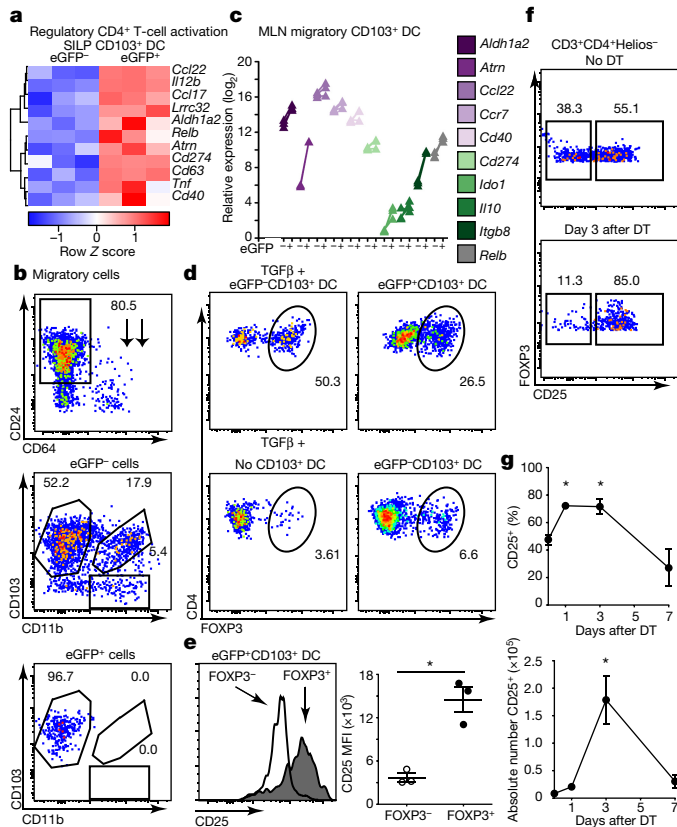


Figure 4 | Apoptotic cell cargo endow migratory CD103⁺ dendritic cells with the ability to induce T_{reg}-cell differentiation. **a**, Hierarchical clustering of differentially expressed genes upregulated in CD103⁺ dendritic cells as in Fig. 3. **b–f**, Flow cytometry of untreated VDTR MLN phagocytes (**b**) and splenic CD4⁺ T cells (**d–f**) stained after co-culture with indicated *ex vivo* dendritic cells (**d, e**) or MLN CD4⁺ T cells stained *ex vivo* (**f**). Numbers indicate percentage of cells within gates. **b**, Migratory phagocytes pre-gated on live CD45⁺MHCII^{hi}CD11c⁺ cells. **c**, qRT-PCR for indicated genes representative of three independent analytical repeats with one biological replicate consisting of two mice. **e**, Geometric mean of CD25 (right panel) from **d**. Data represent two independent experiments, 2–3 replicates per condition. **f**, MLN T_{reg} (pre-gated on CD3⁺CD4⁺Helios⁻) cells from VDTR mice untreated (top) and 3 days after 2 ng g⁻¹ diphtheria toxin treatment (bottom). **g**, Quantification of data from **f** over time. *n* = 4 mice per time point; *n* = 5 mice treated with PBS. One-way ANOVA; **P* < 0.05. Data are mean ± s.e.m.

of apoptotic cell sampling, and point to a possible disruption of the tolerogenic consequences of apoptotic IEC sampling during IBD.

Apoptotic cells mediate their known roles in immunosuppression by pairing downregulation of inflammatory genes with upregulation of negative regulatory nodes that simultaneously target more than one pathway of inflammation. Apoptotic IECs also bestow the ability to instruct T_{reg}-cell differentiation upon the single intestinal CD103⁺ dendritic cell subset tasked with their sampling. Innate immune recognition of apoptotic IECs is thus a critical component of the mechanisms that mediate gut homeostasis^{3,6}.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 7 December 2015; accepted 30 September 2016.

Published online 9 November 2016.

- Poon, I. K., Lucas, C. D., Rossi, A. G. & Ravichandran, K. S. Apoptotic cell clearance: basic biology and therapeutic potential. *Nat. Rev. Immunol.* **14**, 166–180 (2014).
- Green, D. R., Ferguson, T., Zitvogel, L. & Kroemer, G. Immunogenic and tolerogenic cell death. *Nat. Rev. Immunol.* **9**, 353–363 (2009).

- Blander, J. M. Death in the intestinal epithelium—basic biology and implications for inflammatory bowel disease. *FEBS J.* **283**, 2720–2730 (2016).
- Eisenhoffer, G. T. et al. Crowding induces live cell extrusion to maintain homeostatic cell numbers in epithelia. *Nature* **484**, 546–549 (2012).
- Bekiaris, V., Persson, E. K. & Agace, W. W. Intestinal dendritic cells in the regulation of mucosal immunity. *Immunol. Rev.* **260**, 86–101 (2014).
- Gross, M., Salame, T. M. & Jung, S. Guardians of the gut – murine intestinal macrophages and dendritic cells. *Front. Immunol.* **6**, 254 (2015).
- Khor, B., Gardet, A. & Xavier, R. J. Genetics and pathogenesis of inflammatory bowel disease. *Nature* **474**, 307–317 (2011).
- Schlitzer, A. et al. IRF4 transcription factor-dependent CD11b⁺ dendritic cells in human and mouse control mucosal IL-17 cytokine responses. *Immunity* **38**, 970–983 (2013).
- Shrimpton, R. E. et al. CD205 (DEC-205): a recognition receptor for apoptotic and necrotic self. *Mol. Immunol.* **46**, 1229–1239 (2009).
- Kim, G. H., Dayam, R. M., Prashar, A., Terebiznik, M. & Botelho, R. J. PIKfyve inhibition interferes with phagosome and endosome maturation in macrophages. *Traffic* **15**, 1143–1163 (2014).
- Fang, W. F. et al. 5-Lipoxygenase activating protein (FLAP) dependent leukotriene biosynthesis inhibition (MK591) attenuates Lipid A endotoxin-induced inflammation. *PLoS One* **9**, e102622 (2014).
- Blander, J. M. A long-awaited merger of the pathways mediating host defence and programmed cell death. *Nat. Rev. Immunol.* **14**, 601–618 (2014).
- Arbore, G. & Kemper, C. A novel “complement–metabolism–inflammasome axis” as a key regulator of immune cell effector function. *Eur. J. Immunol.* **46**, 1563–1573 (2016).
- Hu, H. & Sun, S. C. Ubiquitin signaling in immune responses. *Cell Res.* **26**, 457–483 (2016).
- Lee, M. S., Kim, B., Oh, G. T. & Kim, Y. J. OASL1 inhibits translation of the type I interferon-regulating transcription factor IRF7. *Nat. Immunol.* **14**, 346–355 (2013).
- Wakioka, T. et al. Spred is a Sprouty-related suppressor of Ras signalling. *Nature* **412**, 647–651 (2001).
- Guo, Z. et al. CD4⁺CD25⁺ regulatory T cells in the small intestinal lamina propria show an effector/memory phenotype. *Int. Immunol.* **20**, 307–315 (2008).
- Tran, D. Q. et al. GARP (LRRC32) is essential for the surface expression of latent TGFβ on platelets and activated FOXP3⁺ regulatory T cells. *Proc. Natl Acad. Sci. USA* **106**, 13445–13450 (2009).
- Josefowicz, S. Z., Lu, L. F. & Rudensky, A. Y. Regulatory T cells: mechanisms of differentiation and function. *Annu. Rev. Immunol.* **30**, 531–564 (2012).
- Maldonado, R. A. & von Andrian, U. H. How tolerogenic dendritic cells induce regulatory T cells. *Adv. Immunol.* **108**, 111–165 (2010).
- Kim, K. S. et al. Dietary antigens limit mucosal immunity by inducing regulatory T cells in the small intestine. *Science* **351**, 858–863 (2016).
- Jostins, L. et al. Host–microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* **491**, 119–124 (2012).
- Moon, C. M. et al. Genetic variants in the IL12B gene are associated with inflammatory bowel diseases in the Korean population. *J. Gastroenterol. Hepatol.* **28**, 1588–1594 (2013).
- Guo, X. et al. Disruption of inducible 6-phosphofructo-2-kinase impairs the suppressive effect of PPARγ activation on diet-induced intestine inflammatory response. *J. Nutr. Biochem.* **24**, 770–775 (2013).
- Haberman, Y. et al. Pediatric Crohn disease patients exhibit specific ileal transcriptome and microbiome signature. *J. Clin. Invest.* **124**, 3617–3633 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements We are grateful to S. V. Chittur and M. Kuentzel, SUNY at the Albany Center for Functional Genomics. We thank M. Bogunovic at Pennsylvania State University, S. Jung at the Weizmann Institute of Science, Blander Laboratory members, J. Ochando and C. Bare at the Icahn School of Medicine Flow Cytometry Core, and M. A. Blander and S. J. Blander for discussions, help, and support. This work was supported by institutional seed funds to J.M.B. J.M.B. and her laboratory were supported by NIH grants AI095245, AI123284, DK072201, the Burroughs Wellcome Fund, and the Leukemia and Lymphoma Society. R.J.C. was supported by NIH training grants 2T32A1007605-11 and 5T32DK007792-12. G.Ba. was supported by the Crohn's and Colitis Foundation of America (CCFA) Research Fellowship Award. B.M.H.: NIAID contract HHSN272201000054C and U19 AI117873. J.C.: R01 DK092235, U01 DK62429, U01 DK062422, philanthropic SUCCESS, Sanford J. Grossman Charitable Trust. S.A.L. and G.C.F.: NIH 5P01DK072201-09 and 5R01CA161373-04, CCFA 330239, and SUCCESS. G.Bo.: Jenna and Paul Segal grant.

Author Contributions R.J.C. and J.M.B. designed the study and wrote the manuscript. R.J.C. conducted most experiments; G.Ba. performed the initial set-up, protocol optimization and provided sorting expertise; L.M. performed initial VDTR characterizations; G.Bo. provided microarray and statistical analysis expertise; B.M.H. performed ImageStream acquisition and analyses; J.C. and K.G. performed IBD GWAS data comparison with differentially expressed phagocyte genes; G.C.F. and S.A.L.: VDTR strain derivation and data discussions. J.M.B. conceived the study.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.M.B. (JMBlander@med.cornell.edu) or S.A.L. (Sergio.lira@mssm.edu).

METHODS

Mice. Plasmids containing the 9-kb mouse villin promoter (pBS-Villin)^{26,27} and simian diphtheria toxin receptor (*HBEGF* ('DTR')) with the enhanced green fluorescent protein (pDTR-eGFP) fusion gene²⁸ have been previously described. The pDTR-eGFP was PCR amplified with primers harbouring a 5' BsiWI site and a 3' MluI site and cloned into pBS-Villin. The pBS-Villin/DTR-eGFP plasmid was verified by sequencing and the transgene was isolated from the plasmid by restriction enzyme digestion and gel purification. The transgene was microinjected into fertilized eggs from C57BL/6J mice (Jackson Laboratory) and transferred into oviducts of ICR foster mothers as previously described²⁶. Identification of the transgenic mice was performed by PCR amplification using the following primers: 5'-ACTGCTCTCACATGCCTTCT-3' and 5'-CTTCTCCCTAGTCCCTTGC-3'. For diphtheria toxin administration, mice were injected intraperitoneally with 2 or 10 ng g⁻¹ diphtheria toxin (EMD Chemicals) and humanely killed 1–24 h later²⁹. Control mice were injected with PBS. For dextran sulphate solution (DSS) (MP Biomedicals) studies, mice were supplemented with 3% DSS in the drinking water for five days. On day three, water bottles were refilled with 3% DSS solution and on day five, replaced with fresh drinking water. Mice were weighed and monitored daily for signs of distress, morbidity or mortality during the course of the experiment until they were killed on day 7. Both male and female mice ages 6–8 weeks were used for all studies. All experiments were approved by the institutional animal care and use committee and carried out in accordance with the 'Guide for the Care and Use of Laboratory Animals' (NIH publication 86–23, revised 1985).

Isolation of professional phagocytes from the SILP. Before isolating professional phagocytes ('phagocytes'), VDTR and VDTR negative littermate controls were intraperitoneally injected with PBS (vehicle) or diphtheria toxin (EMD Chemicals) at a low (2 ng g⁻¹) or high (10 ng g⁻¹) dose per body weight. Mice were then killed 1–24 h later and phagocytes were isolated from the SILP as previously described with some modifications³⁰. In brief, the small intestine, including the duodenum, jejunum and ileum, was excised and Peyer's patches removed. Next, the small intestine was opened longitudinally with surgical scissors and flushed with ice-cold PBS to remove the faecal content. Intestines were then cut into 0.5-cm pieces and transferred into 50-ml conical tubes containing 20 ml of PBS. Samples were then vigorously shaken for 30 s using the vortex genie (Scientific Industries) and passed over 100-µm nylon cell strainers (BD Falcon). Fresh PBS was added to the tissue samples and the shaking and filtering process was repeated a total of eight times. To isolate and remove the intestinal epithelial cell layer, samples were washed with 20 ml of warm PBS containing 3 mM EDTA and passed over cell strainers. This was repeated three times. Flow-through was kept as purified for IECs, whereas whole tissues were further processed to isolate dendritic cell and macrophage subsets. Next, samples were washed with ice-cold PBS followed by RPMI 1640 (Sigma) containing 5% FBS to remove the EDTA. Samples were then re-suspended with RPMI 1640 containing 5% FBS, 1 mg ml⁻¹ collagenase D (Roche), and 1 mg ml⁻¹ DNase I (Roche) and incubated in a 37 °C water bath for 60 min. Samples were shaken every 20 min during this time. At the completion of the incubation, samples were washed with FACS buffer to remove the collagenase and then passed through an 18-gauge needle followed by a 21-gauge needle to create a single-cell suspension. Phagocytes were then enriched from samples by using a 1.065 g ml⁻¹ OptiPrep (Sigma) density gradient according to the manufacturer's protocol. Following centrifugation, phagocytes were isolated from both low- and mid-density bands and finally re-suspended in FACS buffer for flow cytometric analyses.

Flow cytometric analyses. Mouse spleen was digested in parallel with small intestine samples and used for single-colour compensation controls. All samples were pretreated with Fc block for 10 min at 4 °C followed by fluorescently conjugated antibody labelling at 4 °C for 60 min. The following antibodies were used for these studies: Antibodies from BioLegend including Alexa Fluor 647- or 700-conjugated anti-CD11c (clone N418), PerCP/Cy5.5-conjugated anti-CD24 (clone M1/69), APC/Cy7-conjugated anti-CD45 (clone 30-F11), APC-conjugated anti-CD64 (clone X54-5/7.1), APC-conjugated anti-CD274 (clone 10F9G2), PerCP/Cy5.5-conjugated anti-F4/80 (clone BM8), Alexa Fluor 700-conjugated anti-Ly-6c (clone HK1.4), and Phycoerythrin (PE) or Brilliant Violet 421-conjugated anti-MHCII I-A/I-E (clone M5/114.15.2); antibodies from eBioscience including FITC-conjugated anti-CD4 (clone RM4-5), PE/Cy7-conjugated anti-CD11b (clone M1/70) and PE-conjugated anti-CD103 (clone 2E7); and TxRed-conjugated anti-CD45 (clone 30-F11) from Invitrogen. Live/Dead Aqua (Life Technologies) was used to discriminate viable cells.

Identification of SILP phagocytes bearing apoptotic IECs. Phagocytes isolated from the SILP were surface stained for: APC/Cy7-conjugated anti-CD45, Alexa Fluor 700-conjugated anti-CD11c, Brilliant Violet 421-conjugated anti-MHCII I-A/I-E, PE/Cy7-conjugated anti-CD11b, PE-conjugated anti-CD103, PerCP/Cy5.5-conjugated anti-CD24, and APC-conjugated anti-CD64. The identification of phagocytes from VDTR mice with IEC cargo was determined by the presence of eGFP and this gate was defined on the basis of C57BL/6J and VDTR⁻

littermate controls that were eGFP⁻. Sample acquisition was performed using the LSRFortessa (BD Biosciences) and data analyses were performed using the FlowJo analytical software (Tree Star).

Cell sorting. To sort phagocytes with and without apoptotic IEC cargo, the following surface markers were used: APC/Cy7-conjugated anti-CD45, Alexa Fluor 700-conjugated anti-CD11c, Brilliant Violet 421-conjugated anti-MHCII I-A/I-E, PE/Cy7-conjugated anti-CD11b, PE-conjugated anti-CD103, PerCP/Cy5.5-conjugated anti-CD24, and APC-conjugated anti-CD64. The identification of phagocytes from VDTR mice with IEC cargo was determined by the presence of eGFP and this gate was defined on the basis of C57BL/6J and VDTR⁻ littermate controls that were eGFP⁻. Sorted populations were live, CD45⁺MHCII⁺CD11c⁺ phagocytes that were either eGFP⁻ or eGFP⁺ including (i) CD103⁺CD11b⁻CD24⁺CD64⁻ (hereafter CD103), (ii) CD103⁺CD11b⁺CD24⁻CD64⁺ (hereafter CD103 CD11b), and (iii) CD103⁻CD11b⁺CD24⁻CD64⁺ (hereafter CD11b) for a total of six populations. Owing to the four-sample sort-maximum of the instrument, the three eGFP⁺ populations were collected first and then fresh collection tubes were added for the three eGFP⁻ populations. Cells were sorted directly into 0.5 ml TRIzol LS reagent (Life Technologies) for microarray processing (see below). Each sort was performed at 4 h following diphtheria toxin administration and consisted of 3–4 pooled VDTR mice. The following are the cell yield ranges for each subset: 1,000–5,000 eGFP⁺CD103⁺; 3,000–9,000 eGFP⁺CD103⁺CD11b⁺; 10,000–40,000 eGFP⁺CD11b⁺; 4,500–10,000 eGFP⁻CD103⁺; 40,000–80,000 eGFP⁻CD103⁺CD11b⁺; and 30,000–100,000 eGFP⁻CD11b⁺. FACS was conducted on the FACSaria IIu SORP (BD Biosciences). The following are the RNA yield ranges for each subset: 200–2,400 pg eGFP⁺CD103⁺; 200–3,000 pg eGFP⁺CD103⁺CD11b⁺; 600–3,000 pg eGFP⁺CD11b⁺; 200–4,600 pg eGFP⁻CD103⁺; 600–5,000 pg eGFP⁻CD103⁺CD11b⁺; and 450–4,000 pg eGFP⁻CD11b⁺. The purity and identity of each subset was validated as indicated in Extended Data Fig. 5 and according to markers as previously reported³¹.

Image Stream. For analysis of IEC engulfment by CD11c⁺ phagocytes, single-cell suspensions were prepared as described for flow cytometric analyses and acquired using the IS 100 Imaging flow cytometer (Amnis Corp). Phagocytes with eGFP⁺ cargo were identified as those that contained single nuclei and were CD45⁺, CD11c⁺ and MHCII⁺. Data were analysed using IDEAS software (Amnis Corp) and spectrally compensated using a compensation matrix generated from the following single-colour controls; FITC-conjugated CD4, PE-conjugated MHCII, Alexa Fluor647-conjugated CD11c, TxRed-conjugated CD45, and Hoechst stain.

Real-time qRT-PCR. Total RNA was isolated from mouse small intestine using RNeasy mini-kit (Qiagen) and quantified by a spectrophotometer. Reverse transcription was performed with Superscript III (Invitrogen) and cDNA was synthesized using the Mastercycler ep (Eppendorf). Real-time quantitative RT-PCR was conducted in duplicate on a Viia 7 Real-time PCR System (Life Technologies) using TaqMan quantitative PCR Master Mix at a concentration of 1× (Applied Biosystems) or SYBR Green Real-Time PCR Master Mixes for the eGFP and HBEGF ('DTR') transgenes. Samples were normalized to β-actin and relative expression was determined by 2^{-ΔΔC_t} method. Forward (FW) and reverse (RV) primers for SYBR Green include:

5'-AGCAAAGACCCCAACGAGAA-3' eGFP FW, 5'-GGCGGCGGTACGAA-3' eGFP RV; 5'-CACTTTATCCTCCAAGCCACAAG-3' Hbegf FW, 5'-CTTTTCCCGTGCCTCCTCT-3' Hbegf RV.

All probe sequences are in the format: 5' FAM-sequence-BHQ-1 3' and together with forward (FW) and reverse (RV) primer pairs were synthesized by Biosearch Technologies. 5'-AGCCACCCCACTCCTAAGAGGAGG-3' *Actb* probe, 5'-GAAGTCCCTCACCTCCCAA-3' *Actb* FW, 5'-GGCATGGACGCGACCA-3' *Actb* RV; 5'-AAATCGGTGATCCAGGGATTGTTCCA-3' *Acadsh* probe, 5'-CCTCTGGTTTCTCTATGGATGA-3', *Acadsh* FW, 5'-TCCCTCCATATTGTGCTTCAAC-3' *Acadsh* RV; 5'-CGGGACAGGGCAA CTCTTGCAA-3' *Aldh1a2* probe, 5'-GCTTGCAGACTTGGTGGAA-3' *Aldh1a2* FW, 5'-GCTTGCAGGAATGGCTTACC-3' *Aldh1a2* RV; 5'-CCCCCTTTCCTGTGTACTCTGGAC-3' *Alox5ap* probe, 5'-CAACCAGA ACTCGGTAGATGC-3' *Alox5ap* FW, 5'-GAAGGCGCAGGACTTG-3' *Alox5ap* RV; 5'-TGCCTTTAGTGGCCTCATTGTTCC-3' *Atrn* probe, 5'-GGACTCAA TCTACGCACCTCTGAT-3' *Atrn* FW, 5'-GCCGTCTCATTGCCATCTCTT-3' *Atrn* RV; 5'-TTGGCATCAATCTGAGCTGTGGTG-3' *Axl* probe, 5'-GCCCCATCAACTTCGGAAGAAG-3' *Axl* FW, 5'-CCTCTGGC ACCTGTGATATTC-3' *Axl* RV; 5'-AGTGAAGGAGTTCTTCTGGACCTCAA-3' *Ccl22* probe, 5'-CACCCCTCTGCCATCACGTT-3' *Ccl22* FW, 5'-ATCTCGGTTCTTGACGGTTATCA-3' *Ccl22* RV; 5'-CCACTGCTCATGGATATGTGAACAATAGAGACC-3' *Ccr2* probe, 5'-AGGGTCACAGGATTAGG AAGGTT-3' *Ccr2* FW, 5'-CGTTCTGGGCACCTGATTAA-3' *Ccr2* RV; 5'-CAG TGCCCAAGTGGAGGCTTGATC-3' *Ccr7* probe, 5'-CAGCGTG AGATGCTCACTAGG-3' *Ccr7* FW, 5'-ATCTGGGCCACTTGATGG-3'

Ccr7 RV; 5'-AGATTTCGCTGTACCAGCACAGACA-3' *Cd40* probe, 5'-TC TCAGCCCAGTGGAAACA-3' *Cd40* FW, 5'-CGGTGCCCTCCTTCTTAACC-3' *Cd40* RV; 5'-CGAATCACGCTGAAAGTCAATGCC-3' *Cd274* probe, 5'-CG GTGGTGCGGACTACAAG-3' *Cd274* FW, 5'-CCCTCGGCCTG ACATATTAGTTC-3' *Cd274* RV; 5'-TTCCCAGGGCTTGAGGCTCCC-3' *Cd300a* probe, 5'-GGCCACCGTGAACATGACTA-3' *Cd300a* FW, 5'-GCAG GAGAGCTAACACAGACAGAAC-3' *Cd300a* RV; 5'-ATGGAATAATGG GTGGCGCTAACCCA-3' *Cfh* probe, 5'-CCGAACACTTGGCACTATTGTAA-3' *Cfh* FW, 5'-CTCCGGGATGCCACAAG-3' *Cfh* RV; 5'-CCCTGAAC AACCAACAGATGACACTGG-3' *Elf3* probe, 5'-GGCACTGAAGAC TTGGTGTG-3' *Elf3* FW, 5'-CCCTGAACAACCAACAGATGACACTGG-3' *Elf3* RV; 5'-AGCTGACAGATACACTCCAAAGCGGA-3' *Fos* probe, 5'-AG TGGCGAATCGGAGGA-3' *Fos* FW, 5'-TGCAACGCAGACTTCTCATC-3' *Fos* RV; 5'-CTGCTCCTGCTGGCTTCCGAGT-3' *Gas6* probe, 5'-CTGGGCACTGCGCTTCTG-3' *Gas6* FW, 5'-CGCAACAGCACAGTGTGA-3' *Gas6* RV; 5'-TCTTATGCAGACTGTGTCTGGCA-3' *Ido1* probe, 5'-GGGCC TGCCTCTATTCTG-3' *Ido1* FW, 5'-CCCACCAGGAAATGAGAACAGA-3' *Ido1* RV; 5'-TCACAAGCAGACAGACAAGCCTGTC-3' *Il1b* probe, 5'-AAAGAC GGCACACCCACCCTGC-3' *Il1b* FW, 5'-TGTCTGACCACTG TTGTTCCAG-3' *Il1b* RV; 5'-TCTGCAAGAGACTTCCATCCAGTTGCCT-3' *Il6* probe, 5'-CCAGAAACCGCTATGAAGTTCC-3' *Il6* FW, 5'-TCACCA GCATCAGTCCCAAG-3' *Il6* RV; 5'-TTCAAACAAGGACCACTGGACA-3' *Il10* probe, 5'-TCAGCCAGGTGAAGACTTTC-3' *Il10* FW, 5'-GGCAA CCAAGTAACCCCTTA-3' *Il10* RV; 5'-TAACTGGGATCCAGGCACGCC-3' *Ly75* probe, 5'-GTCAGACTTACGGCCACTCAA-3' *Ly75* FW, 5'-TGACCCA CCAATCACAGGT-3' *Ly75* RV; 5'-TCCCTTACTTTATTAAGCAG CTTGAGAGTG-3' *Mertk* probe, 5'-TGATCCCATATACGTGGAAGTTCA-3' *Mertk* FW, 5'-CCTGGCAGGTGAGTTGAAG-3' *Mertk* RV; 5'-TTTGC GTCTGACTGCCGAGACTC-3' *Muc2* probe, 5'-CCTGGCCT CTGTGATTACAAC-3' *Muc2* FW, 5'-GGTGCACAGCAATTCCTGTAG-3' *Muc2* RV; 5'-TCGCAACCAGATCGGAGATGTGG-3' *Nlr5* probe, 5'-CCA GAACCTCAGGAAATTTGACTGA-3' *Nlr5* FW, 5'-TTTGCGCAAGA TGGCAGCTAA-3' *Nlr5* RV; 5'-CTGCTGCCTCACTTCTAGCTTCTGC-3' *Nlrp3* probe, 5'-GTTGCCTGTTCTTCCAGACT-3' *Nlrp3* FW, 5'-GGC TCCGGTTGGTGCTTAG-3' *Nlrp3* RV; 5'-TAGGCTGCTTTGGGAA TGGCACC-3' *Oasl1* probe, 5'-CGCGTGCTCAAGGTACTAAG-3' *Oasl1* FW, 5'-GACCAGCTCCACGTCTGTAG-3' *Oasl1* RV; 5'-TTGTGATGACTACATG GTCACACTCTTC-3' *Plac8* probe, 5'-GAACCCGATACGGCATCTCT-3' *Plac8* FW, 5'-TCTTGCCATCCAGCTCCTTAG-3' *Plac8* RV; 5'-ACCAACACATCG GAGCTGCGGA-3' *Relb* probe, 5'-GAGCCTGTCTACGACAAGAAGTC-3' *Relb* FW, 5'-GCCCCGCTCCTTGTGATTTC-3' *Relb* RV; 5'-AGTTATGCACGAGT GCGAGCTGT-3' *Spre1* probe, 5'-CGGCGACTTCTGACAACGATA-3' *Spre1* FW, 5'-GGTAGCCATCCACCACTGTAG-3' *Spre1* RV; 5'-AGAGGTG ACCCGCGTGCTAATGGTG-3' *Tgfb1* probe, 5'-CCCGAAGCGGACT ACTATGC-3' *Tgfb1* FW, 5'-ATAGATGGCGTTGTTGCGGT-3' *Tgfb1* RV; 5'-CTCTGCCTGCATCCAATCACTCTCA-3' *Timd4* probe, 5'-GGTCCG CCTTCACTACAGAAC-3' *Timd4* FW, 5'-GGCCTGAGTACGGCTATGTC-3' *Timd4* RV; 5'-TGGGCTTCCGAATTCACCTGGAGC-3' *Tnf* probe, 5'-ATGC ACCACCATCAAGGACTCAA-3' *Tnf* FW, 5'-ACCACTCTCCCT TTGCAGAACTC-3' *Tnf* RV; 5'-TCAACTGGTGTCTGTAAGTCAGGA-3' *Tnfai3* probe, 5'-TCCCTGGAAGCCAGAAGAAG-3' *Tnfai3* FW, 5'-GA GGCAGTTTCCATCACCATTG-3' *Tnfai3* RV; 5'-TCCGGAGCTAC TTCAAGCAAGGC-3' *Vill* probe, 5'-GGCAACGAGAGCGAGACTT-3' *Vill* FW, 5'-CGCTGGACATCACAGGAGTT-3' *Vill* RV.

cDNA Microarrays. A total of five sorting experiments with a pool of 3–4 mice were performed for the cDNA microarrays. Following cell sorting into TRIzol LS reagent, samples were shipped on dry ice to the Center for Functional Genomics and the Microarray & HT Sequencing Core Facility at the University at Albany (Rensselaer). A sample clean-up step was performed using RNeasy columns (Qiagen) that included DNase treatment. The isolated RNA was checked for quality using NanoDrop (Thermo Scientific) and Bioanalyzer (Agilent), following which 1 ng of total RNA was processed using WT-Ovation Pico RNA Amplification System (NuGEN). A total of three biological replicates were used for the microarray. When required, RNA was pooled from additional sorts to achieve the 1 ng of total RNA needed for the amplification system. The following are the sort experiments used for each sample: (2, 2 and 5) eGFP⁺CD103⁺; (2, 3, 2 and 5) eGFP⁺CD103⁺CD11b⁺; (2, 4, 5) eGFP⁺CD11b⁺; (3, 2 and 4 and 5, 2 and 4 and 5) eGFP⁺CD103⁺; (2, 3, 4) eGFP⁺CD103⁺CD11b⁺; (2, 3, 5) eGFP⁺CD11b⁺. RNA was reverse-transcribed and sense-target cDNAs were targeted using the standard NuGEN protocol and hybridized to Affymetrix mouse Gene 2.0 ST arrays. These arrays were then washed, stained on a FS 450 station, and scanned on a GeneChip 3000 7G scanner using Affymetrix GeneChip Command

Console Software (AGCC). The Affymetrix microarray data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus (GEO)³² and are accessible with the GEO series accession number GSE85682.

Code availability. Fold changes and statistical significance were identified as those genes that were differentially expressed between eGFP⁺ and eGFP[−] subsets by at least 1.2 fold (ANOVA (Benjamini–Hochberg false discovery rate correction $Q < 0.05$) and Tukey's HSD post-hoc test ($P < 0.05$; $-1.2 > \text{fold} > 1.2$)) and determined using R software (version 3.2.0). Hierarchical clustering of differentially expressed genes meeting the aforementioned criteria were Z-scored and plotted with heatmap.2 (gplots version 2.17.0, CRAN/R). Principal component analyses of the 1,534 genes (ANOVA (Benjamini–Hochberg false discovery rate correction $Q < 0.05$); 4.8% of total) with the most variable expression in each CD11c⁺ subset with and without eGFP cargo were generated using R software which are freely available online.

Microscopy. Paraffin sections. Small and large intestine were dissected and fixed in 10% formalin (Fisher Scientific) for 24 h and then processed for paraffin embedding. Tissue blocks were then cut into 5-mm sections, de-paraffinized by xylene immersion, and hydrated by serial immersion in 100%, 90%, 80%, 70% ethanol and PBS. Antigen retrieval was performed by heating samples in a pressure cooker (Cuisinart) in citrate buffer solution (10 mM citric acid monohydrate, 0.05% Tween 20 and PBS). Sections were then washed twice in PBS, blocked for 30 min in blocking buffer (10% BSA, 0.3% Triton X-100 (Sigma) and TBS), and prepared for labelling. TdT-mediated dUTP nick end labelling (TUNEL) was performed using the *in situ* cell-death detection kit, TMR red (Roche), per the manufacturer's instructions, stained with DAPI, and mounted using Fluoromount-G (Southern Biotech). For cleaved caspase-3 (Cell Signaling), samples were labelled for 60 min at room temperature, stained with DAPI, and mounted using Fluoromount-G. For paraffin images, eGFP signal was not present owing to sample quenching following paraffin embedding and processing.

Frozen sections. Small and large intestine were dissected and fixed overnight in 1.6% paraformaldehyde (Thermo Scientific) containing 20% sucrose at 4 °C. Samples were then placed in OCT (Tissue-Tek) and snap-frozen over dry ice. Tissue sections of 8-mm thickness were cut, air-dried and blocked using blocking solution. Tissues were then labelled using an Alexa Fluor 594-conjugated phalloidin (Invitrogen) or a primary mouse anti-mouse pan-cytokeratin antibody (clone PCK-26) (Abcam) for 60 min in a humidified atmosphere followed by a secondary goat anti-mouse Alexa Fluor 594 (Thermo Fisher Scientific) for 30 min, then stained with DAPI, and mounted using Fluoromount-G. For fluorescent *in situ* hybridization, small intestine and large intestine were dissected and prepared as described for frozen sections³³. Following tissue blocking, sections were incubated with 0.45 pmol μl^{−1} eubacterial oligonucleotide probe (AminoC6 + Alexa Fluor 594) 5'-GCTGCCTCCCGTAGGAGT-3'; (Operon)³³ in a pre-chilled hybridization buffer (Sigma) overnight at 4 °C. Sections were counterstained with DAPI and mounted with Fluoromount-G.

Whole-mount staining. To label small intestine tissues, the whole-mount histology protocol was modified from previously described methods³⁴. In brief, small intestine samples were excised, opened longitudinally, and washed in ice-cold PBS. Samples were then cut to 1 cm in length and placed in 6-ml polypropylene tubes (BD Biosciences). Next, samples were incubated with Fc block at 10 μg ml^{−1} in 200 μl of 2% paraformaldehyde with 1% FBS, 0.3% Triton X-100 in PBS for 3 h at 4 °C with gentle rocking. After blocking and fixing, samples were put into new polypropylene tubes and labelled using 3 μg ml^{−1} of the following antibodies: PE-conjugated anti-CD11c (clone N418) (eBioscience), APC-conjugated anti-CD31 (clone 390) (eBioscience) and anti-cleaved caspase-3 at 1:100. All labelling was conducted in the dark at 4 °C with gentle rocking for 3 h. Finally, samples were washed for 30 min in the dark at 4 °C with fresh PBS and mounted for imaging.

Conventional microscopy was performed using the Eclipse Ni-E motorized upright microscope (Nikon) and images were acquired from paraffin, frozen, and whole mount tissue sections using a Nikon DS-Qi1 Mc camera. Cell quantification was calculated using NIS Elements imaging software (Nikon) and the object count application including intensity of stain thresholds and area restriction filters. Confocal microscopy was performed at the Microscope CORE at the Icahn School of Medicine at Mount Sinai using the Leica SP5 DM upright microscope and Leica LAS AF software.

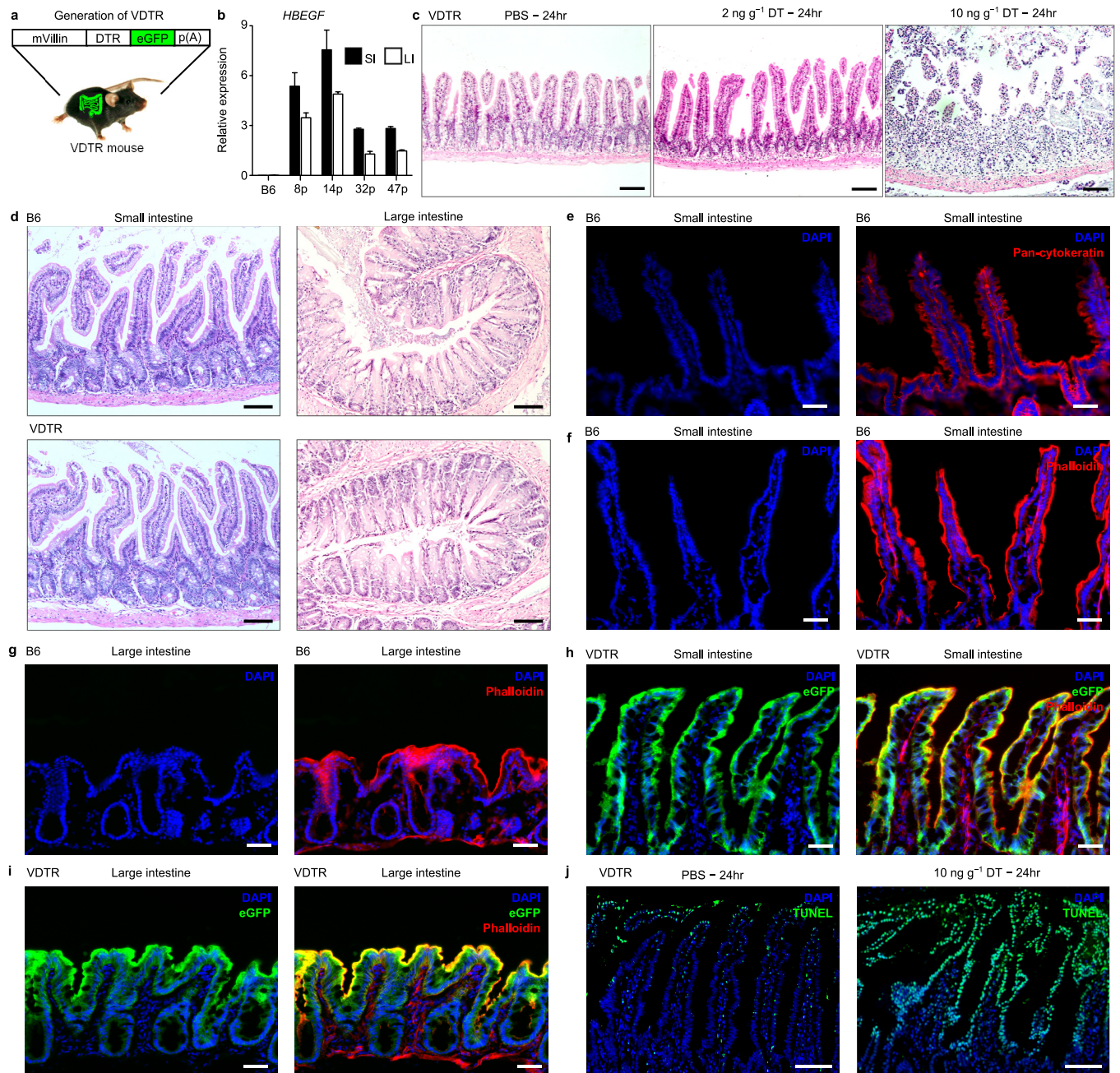
In vitro T-cell differentiation. Naive mouse splenic CD4⁺ T cells were isolated by sorting with MACS CD4⁺ beads (Miltenyi Biotech) according to the manufacturer's instructions and then by FACS using the FACSaria IIu SORP. T cells were sorted on the basis of the following criteria: live, CD45⁺CD3⁺CD4⁺CD25[−]CD44^{−/low}CD62L^{+/high}. Surface antibodies for sorting included: APC/Cy7-conjugated anti-CD45, eFluor 450-conjugated anti-CD3 (clone 145-2c11), PE-conjugated anti-CD4, APC-conjugated anti-CD25 (clone PC61.5),

FITC-conjugated CD62L (clone MEL-14), and Alexa 700-conjugated anti-CD44 (clone IM7) (all eBiosciences). 1×10^5 T cells were then cultured with 1×10^4 eGFP⁺ or eGFP⁻ CD103 dendritic cells sorted from the MLN which were identified as: live, CD45⁺MHCII^{hi}CD11c⁺, eGFP⁺ or eGFP⁻, CD103⁺CD11b⁻ using the aforementioned antibodies for flow cytometry. These cells were cultured in round-bottom 96-well plates (Falcon) with complete IMDM (Gibco) supplemented with 10% FBS, $100 \mu\text{g ml}^{-1}$ penicillin, $100 \mu\text{g ml}^{-1}$ streptomycin, 2 mM L-glutamine, 10 mM HEPES and 1 nM sodium pyruvate for 5 days. Additionally, all cultures were supplemented with $1 \mu\text{g ml}^{-1}$ of soluble anti-CD3 (clone 2C11) as well as 5 ng ml^{-1} of recombinant human anti-IL-2 (Pepro Tech) on days 2 and 4. A total of 2 ng ml^{-1} of recombinant human anti-TGF β 1 (clone 1D11 R&D systems) was added to culture wells where indicated on days 1 and 4. On day 5, cells were first surface stained with FITC-conjugated anti-CD25, PE/Cy7-conjugated anti-CD4, Alexa Fluor 700-conjugated anti-CD3, and APC/Cy7-conjugated anti-CD45, followed by fixation and permeabilization (using the concentrate and diluent provided by eBioscience), and finally intracellular staining for eFluor 450-conjugated anti-FOXP3 (clone FJK-16s), PE-conjugated anti-ROR γ (t) (clone B2D), PerCP-eFluor 710-conjugated anti-GATA-3 (clone TWAJ), and APC-conjugated anti-T-bet (clone eBio4B10) (all eBioscience).

Statistical analysis. Data are presented as mean \pm s.e.m. Statistical significances were determined by a one-way ANOVA with Dunnett's and Newman-Keuls post-tests or unpaired two-tailed *t*-test with Welch correction where specified. *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$. NS, not statistically significant ($P > 0.05$). No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

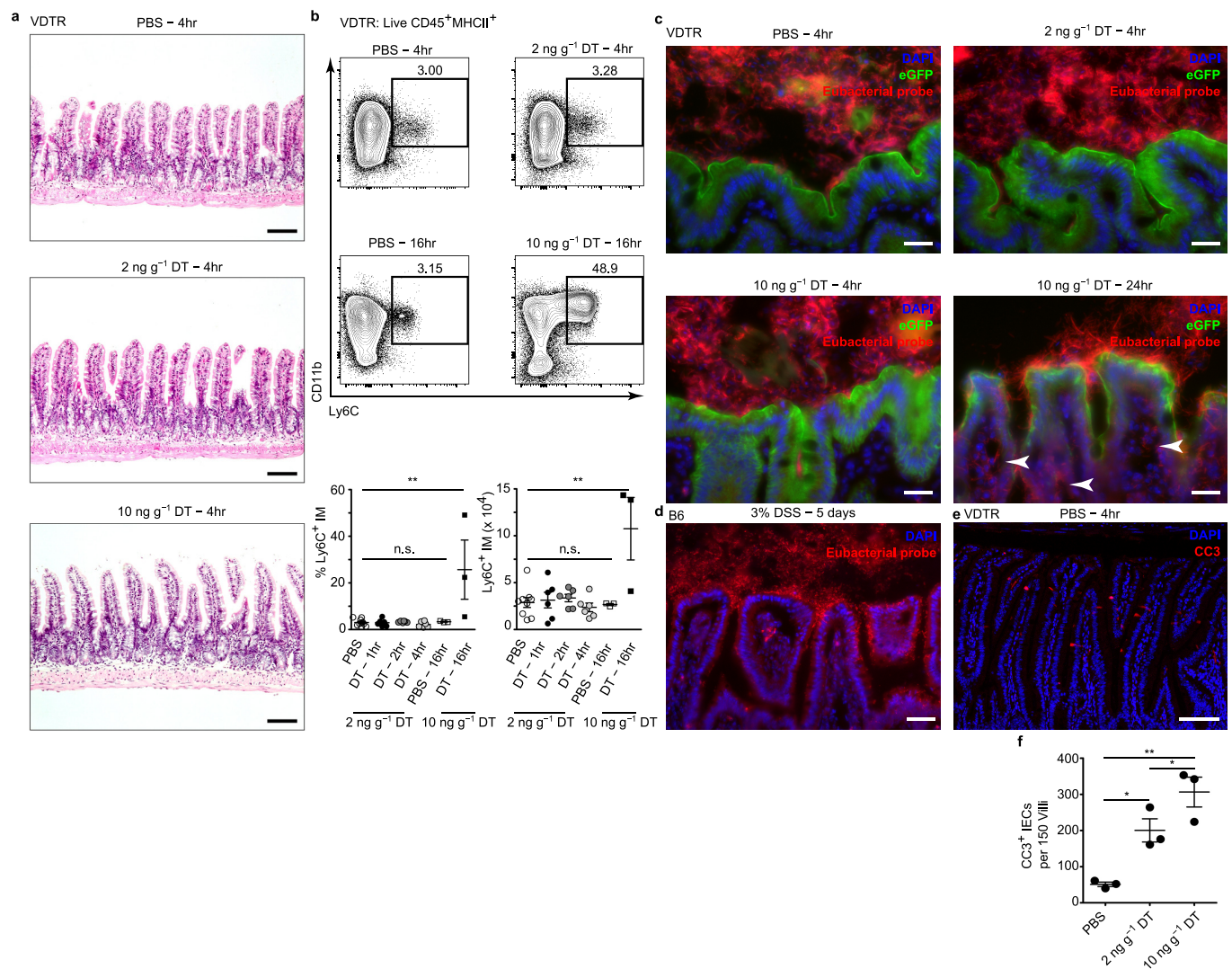
Data availability. The Affymetrix microarray data have been deposited in the NCBI Gene Expression Omnibus (GEO) under GEO series accession number GSE85682.

26. Bongers, G. *et al.* The cytomegalovirus-encoded chemokine receptor US28 promotes intestinal neoplasia in transgenic mice. *J. Clin. Invest.* **120**, 3969–3978 (2010).
27. Madison, B. B. *et al.* Cis elements of the villin gene control expression in restricted domains of the vertical (crypt) and horizontal (duodenum, cecum) axes of the intestine. *J. Biol. Chem.* **277**, 33275–33283 (2002).
28. Jung, S. *et al.* In vivo depletion of CD11c⁺ dendritic cells abrogates priming of CD8⁺ T cells by exogenous cell-associated antigens. *Immunity* **17**, 211–220 (2002).
29. Saito, M. *et al.* Diphtheria toxin receptor-mediated conditional and targeted cell ablation in transgenic mice. *Nat. Biotechnol.* **19**, 746–750 (2001).
30. Bogunovic, M. *et al.* Origin of the lamina propria dendritic cell network. *Immunity* **31**, 513–525 (2009).
31. Miller, J. C. *et al.* Deciphering the transcriptional network of the dendritic cell lineage. *Nat. Immunol.* **13**, 888–899 (2012).
32. Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **30**, 207–210 (2002).
33. Bongers, G. *et al.* Interplay of host microbiota, genetic perturbations, and inflammation promotes local development of intestinal neoplasms in mice. *J. Exp. Med.* **211**, 457–472 (2014).
34. Cummings, R. J. *et al.* Exposure to ionizing radiation induces the migration of cutaneous dendritic cells by a CCR7-dependent mechanism. *J. Immunol.* **189**, 4247–4257 (2012).
35. Shi, C. & Pamer, E. G. Monocyte recruitment during infection and inflammation. *Nat. Rev. Immunol.* **11**, 762–774 (2011).



Extended Data Figure 1 | No gross anatomical changes in the small or large intestine of VDTR mice. **a**, The primate diphtheria toxin receptor (DTR)-eGFP transgene driven by the mouse villin promoter (mVillin). **b**, qRT-PCR for relative expression of primate diphtheria toxin receptor (*HBEGF* ('DTR')) mRNA in both the small and large intestine of four different DTR-eGFP transgenic (VDTR) founder lines. Founder 32p was selected to propagate the VDTR line on the basis of specific transgene expression in the intestinal tract and not in other organs.

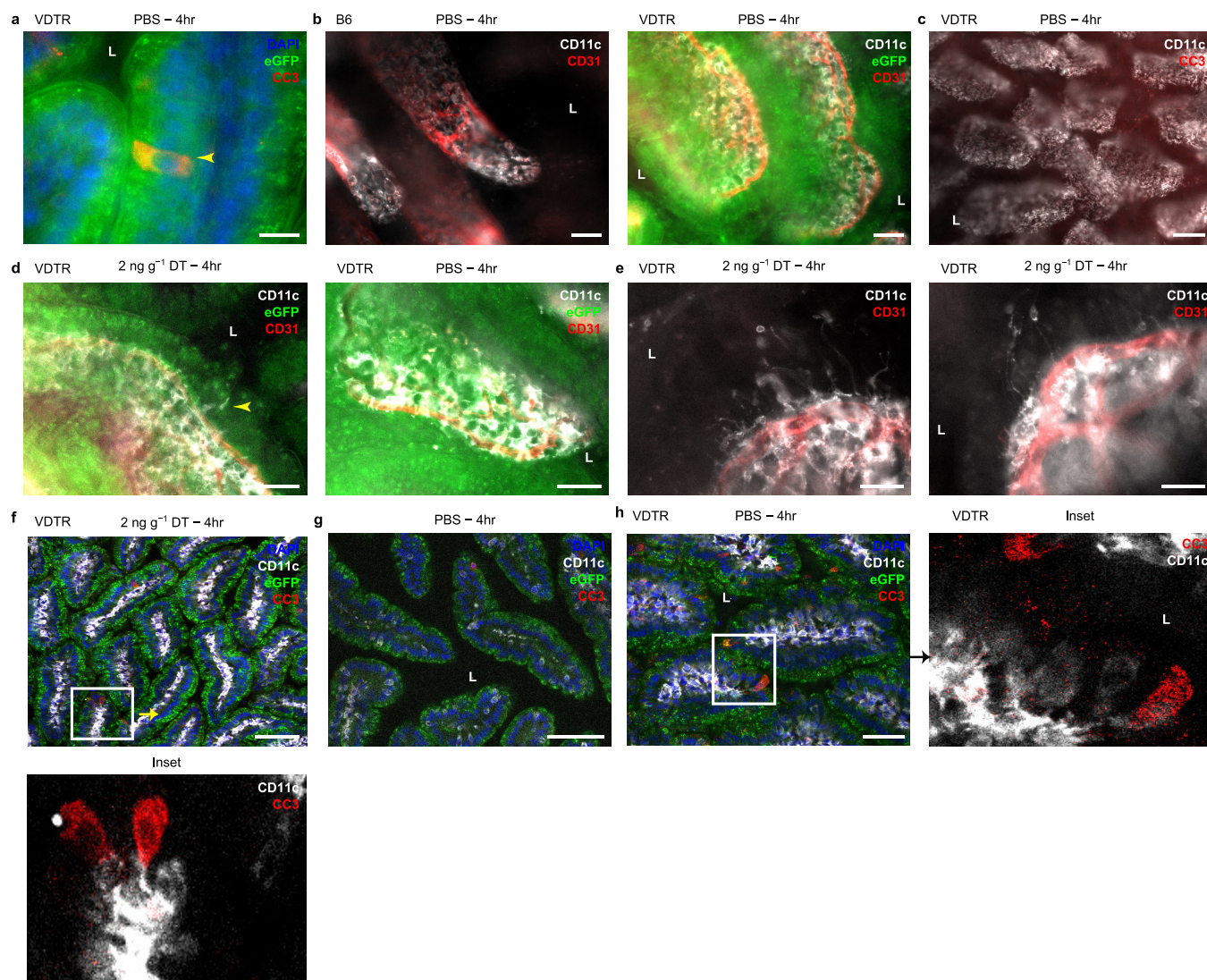
c, d, Haematoxylin and eosin staining (H&E) of small intestine and large intestine paraffin sections from treated VDTR (**c**) and untreated C57BL/6 (B6) and VDTR (**d**) as indicated. **e–i**, Immunofluorescence of small intestine (**e, f, h**) and large intestine (**g, i**) cryo-sections from indicated mice after staining with phalloidin or pan-cytokeratin and DAPI as indicated. **j**, TUNEL (eGFP-quenched after paraffin embedding). p(A), polyadenylation. Scale bars, 100 μ m (**c, d, j**), 50 μ m (**e–i**).



Extended Data Figure 2 | The intestinal microenvironment does not become inflamed following 2 ng g⁻¹ diphtheria toxin administration.

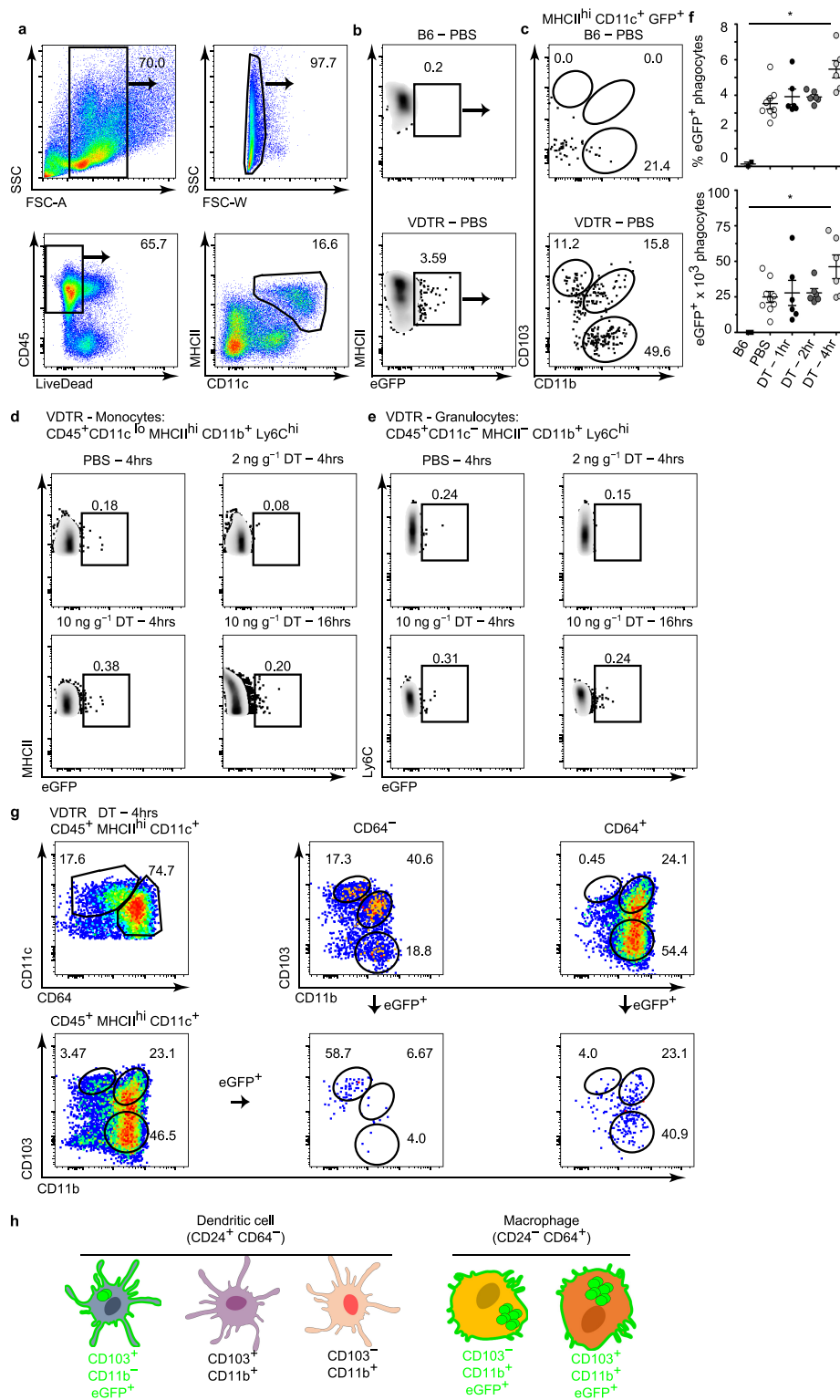
a, Haematoxylin and eosin stain of paraffin-embedded sections from small intestine of VDTR mice at 4 h following PBS or diphtheria toxin administration. **b**, Flow cytometry of SILP cells from VDTR mice treated with PBS or diphtheria toxin. Pre-gated on live CD45⁺MHCII⁺ cells. Numbers above gates indicate percentage of positively stained cells. Quantification is shown in bottom panels. IM, inflammatory monocytes. We noted no tissue destruction or discernible inflammation (**a**), nor infiltration of Ly6C^{hi} monocytes³⁵ (**b**) at 4 h with either dose of diphtheria toxin, in contrast to infiltration of these cells at 16 h following 10 ng g⁻¹ diphtheria toxin. $n = 9$ mice for PBS, $n = 6$ mice for diphtheria toxin at 1–4 h, and $n = 3$ for PBS and diphtheria toxin at 16 h. One-way ANOVA;

$**P < 0.01$; NS, not significant. Flow cytometry gates are representative of at least three independent experiments. Data are mean \pm s.e.m. **c**, **d**, *In situ* hybridization with a eubacterial probe on large intestine cryo-sections following PBS or diphtheria toxin administration to VDTR mice (**c**), or 3% dextran sodium sulphate (DSS) treatment of B6 mice for 5 days (**d**). **c**, Arrowheads indicate presence of eubacterial probe in SILP. There was no bacterial translocation to the intestinal lamina propria at 4-h post-treatment with either 2 ng g⁻¹ or 10 ng g⁻¹ diphtheria toxin, as evidenced by luminal confinement of the *in-situ*-hybridized eubacterial probe signal, and similar to that in PBS-treated controls. **e**, Immunofluorescence on small-intestine paraffin sections stained with antibodies to CC3 at 4 h following PBS injection. **f**, Quantification from **e** and Fig. 1c. Scale bars, 100 μ m (**a**, **e**), 25 μ m (**c**, **d**).



Extended Data Figure 3 | Ileum lamina propria CD11c⁺ phagocytes extend dendrites towards apoptotic intestinal epithelial cells.
a-h, Conventional (a-e) and confocal (f-h) whole-mount microscopy of excised ileum from B6 or VDTR mice following PBS or diphtheria

toxin as indicated. **b, c, e**, eGFP not overlaid. Insets from **f** and **h** depicted in bottom and right panels, respectively, without eGFP overlay. L, lumen. Arrowhead in **a** points to a CC3⁺ IEC; in **d**, to a CD11c⁺ dendrite. Scale bars, 25 μ m (**a, e**), 50 μ m (**b, d, h**), 100 μ m (**c, f, g**).

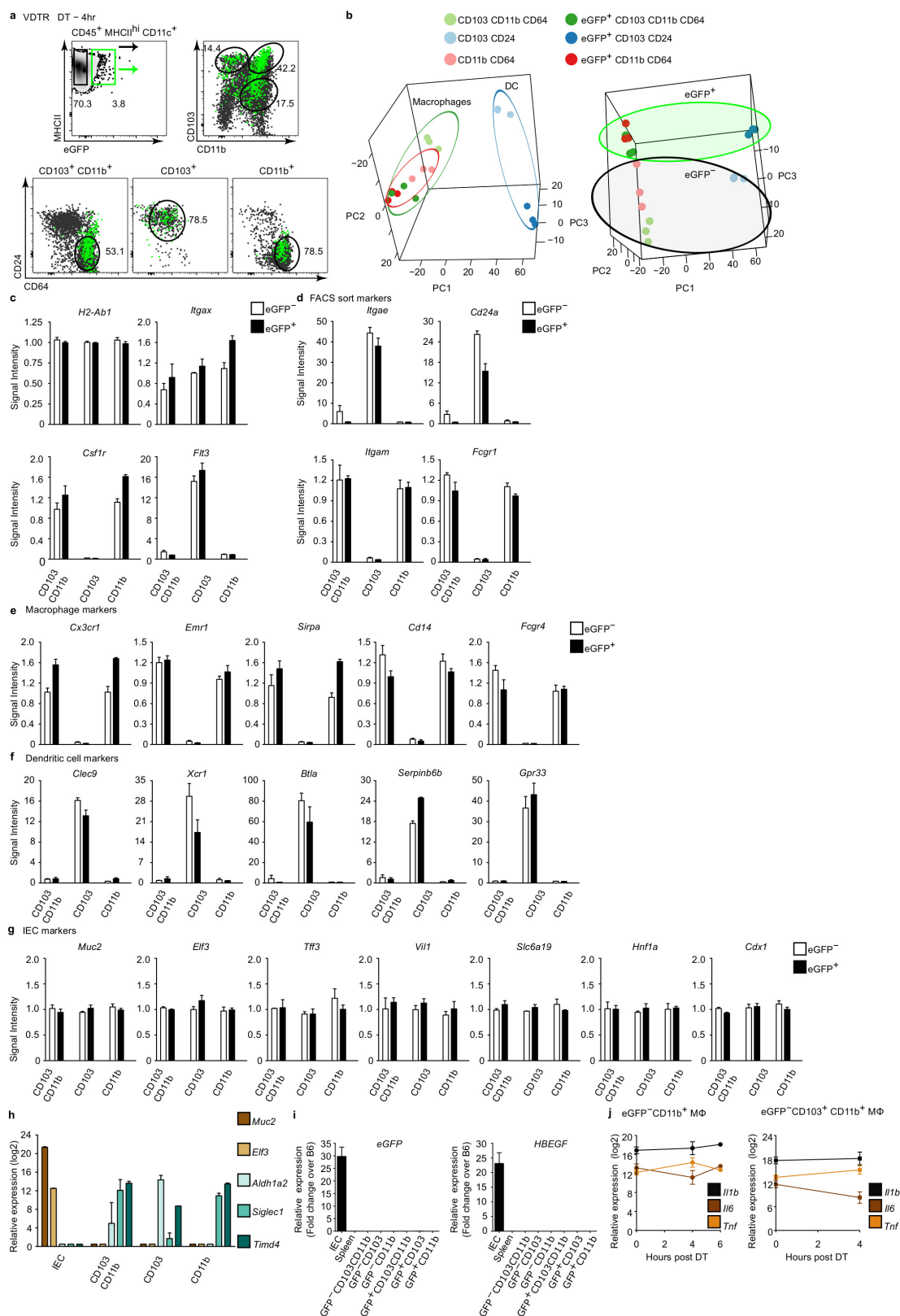


Extended Data Figure 4 | See next page for caption.

Extended Data Figure 4 | Flow cytometry gating strategy for identifying CD11c⁺ and CD11c^{-lo} phagocytes in the SILP of VDTR mice.

a–e, g, Flow cytometric analyses of SILP cells from VDTR mice. Numbers indicate the percentage of gated populations. **a,** Gating strategy for total CD11c⁺ phagocytes. **b,** Identification of the eGFP⁺ gate on the basis of the small intestine cellular profile from B6 (eGFP⁻) mice. **c,** CD103 and CD11b expression on gated cells in **b**. **d,** Monocytes pre-gated on live CD45⁺CD11c^{lo}MHCII^{hi}CD11b⁺Ly6C^{hi} cells. **e,** Granulocytes pre-gated on live CD45⁺CD11c⁻MHCII⁻CD11b⁺Ly6C^{hi} cells. **f,** Percentage of cells and

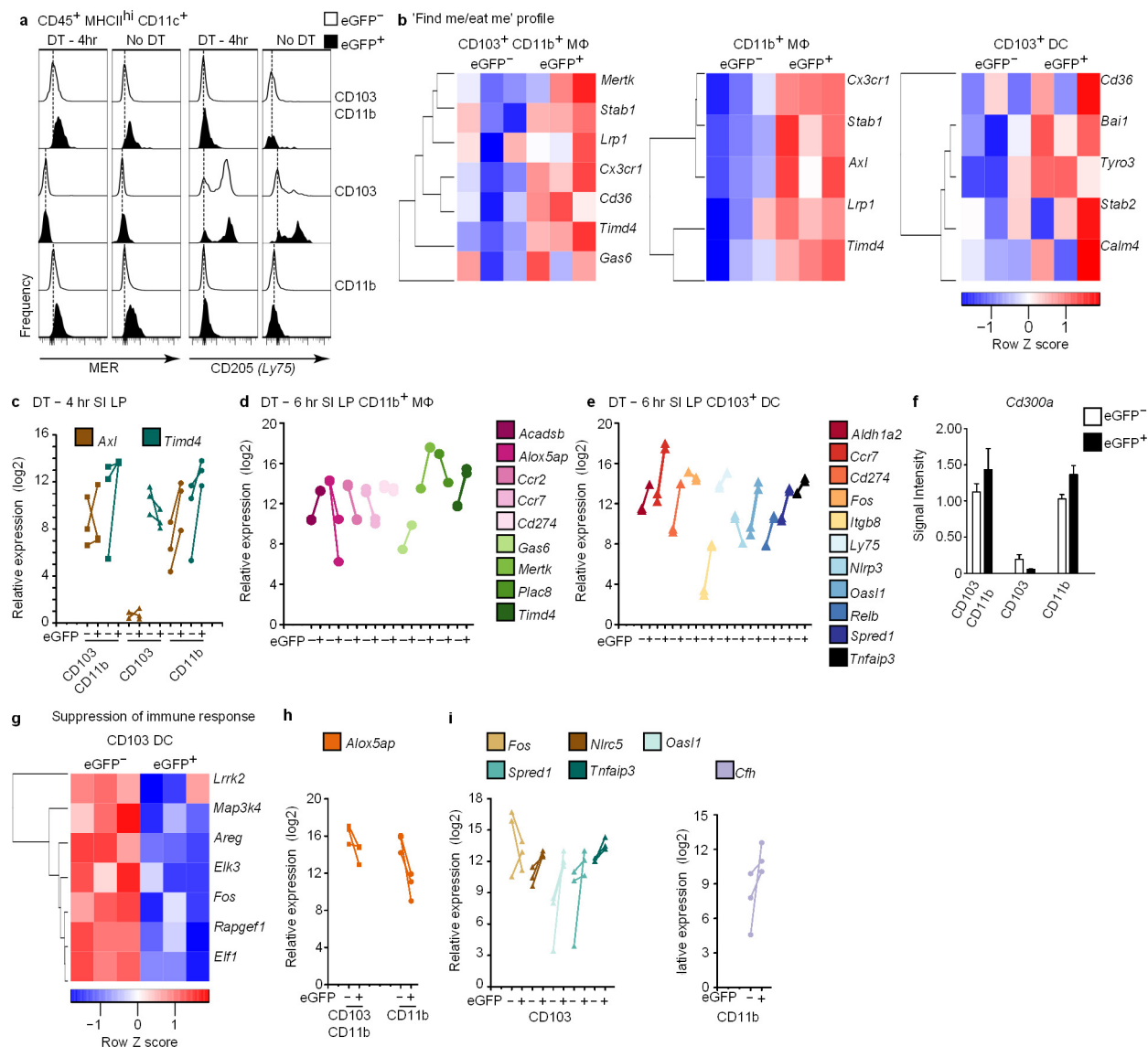
absolute numbers from **b, g**. Dendritic cells and macrophages pre-gated on live CD45⁺MHCII^{hi}CD11c⁺ cells, further identified as CD64⁻ and CD64⁺, respectively. Gating on differential CD64 expression, followed by delineation of CD103 and CD11b expression, also distinguished two macrophage and three dendritic cell populations⁵. Data represent at least three independent experiments and in **f**, $n = 3$ B6 mice; $n = 6$ VDTR mice treated with 2 ng g⁻¹ diphtheria toxin for 1–4 h; $n = 9$ VDTR mice treated with PBS; one-way ANOVA; * $P < 0.05$. Data are mean \pm s.e.m. **h,** Schematic of SILP phagocytes that sample apoptotic IECs (green).



Extended Data Figure 5 | See next page for caption.

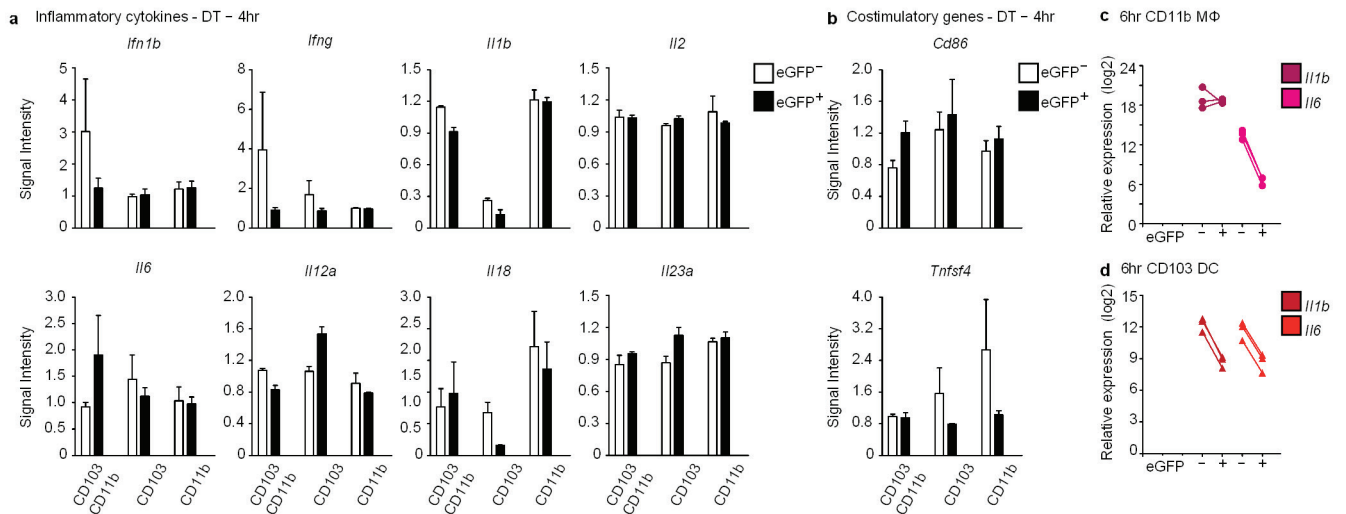
Extended Data Figure 5 | eGFP⁺ and eGFP⁻CD11c⁺ phagocyte populations exhibit distinct transcriptional profiles. **a**, FACS-sorted dendritic cell and macrophage populations from the SILP of VDTR mice at 4 h following diphtheria toxin administration using the gating strategy described in Fig. 2f, which included CD24 and CD64. **b**, Principal Component Analysis (PCA) of the 1,534 genes (ANOVA $Q < 0.05$; 4.8% of total) with most variable expression in dendritic cell and macrophage (M ϕ) subsets, with (eGFP⁺) or without (eGFP⁻) apoptotic IEC cargo. PCA-separated macrophages from dendritic cells (left panel) and eGFP⁺ from eGFP⁻ (right panel). We confirmed the purity and identity of sorted phagocytes on the basis of relative expression of genes encoding the molecules used for FACS sorting, as well as macrophage and dendritic-cell-specific genes. **c–g**, Signal intensity of genes encoding MHCII (*H2-Ab1*) and CD11c (*Itgax*), shared by both dendritic cells and macrophages, the macrophage-specific Csf1 receptor (*Csf1r*), and the dendritic-cell-specific kinase Flt3 (*Flt3*) in CD11c⁺ phagocytes³¹ (**c**); molecules used for FACS sorting (**d**); macrophage markers (**e**); dendritic cell markers (**f**); and IEC markers from analyses of microarray experiments (**g**). It should be noted that IEC-specific genes were not increased in eGFP⁺ compared to eGFP⁻ phagocytes indicating negligible contamination by IEC-specific transcripts. Data represent five independent experiments with four mice per experiment and three biological replicates. White and black bars indicate expression from eGFP⁻ and eGFP⁺ phagocytes, respectively. **h**, Validation by qRT-PCR for IEC-specific genes including *Muc2* (mucin-2) and *Elf3* (E74-like

factor 3), which were not expressed by phagocytes but readily detectable in sorted IECs. Conversely, phagocyte-specific transcripts (*Aldh1a2*, *Siglec1* and *Timd4*)³¹ were expressed in sorted phagocytes but not in IECs. Data represent two independent experiments depicting relative expression (\log_2). **i**, qRT-PCR quantification of relative eGFP expression and the primate diphtheria toxin receptor (*HBEGF* or 'DTR') from VDTR IECs, spleen cells, and eGFP⁻ and eGFP⁺ phagocytes. As expected, expression of the *eGFP* and *DTR* (*HBEGF*) transgenes were confined to IECs in VDTR mice and not present in sorted dendritic cell or macrophages. Data represent two independent experiments depicting fold change over C57BL/6J litter mate controls. **j**, qRT-PCR relative expression (\log_2) of inflammatory transcripts from indicated eGFP⁻ SILP macrophages over time. The unchanged expression of *Il1b*, *Il6*, and *Tnf* in eGFP⁻CD11b⁺ and CD103⁺CD11b⁺ macrophages, which served as sentinels of the microenvironment, provided further evidence for lack of inflammation induction by low-dose 2 ng g⁻¹ diphtheria toxin. Note that the unique transcriptional profiles of eGFP⁺ phagocytes may belong to pre-existing subpopulations that become detectable only when marked by eGFP, owing to an inherently superior capacity for apoptotic cell sampling. However, we consider it more likely that the differential gene expression patterns between eGFP⁺ and eGFP⁻ populations are a consequence of apoptotic IEC internalization. Data represent three independent analytical repeats and one biological replicate consisting of three mice per time point. Data are mean \pm s.e.m.



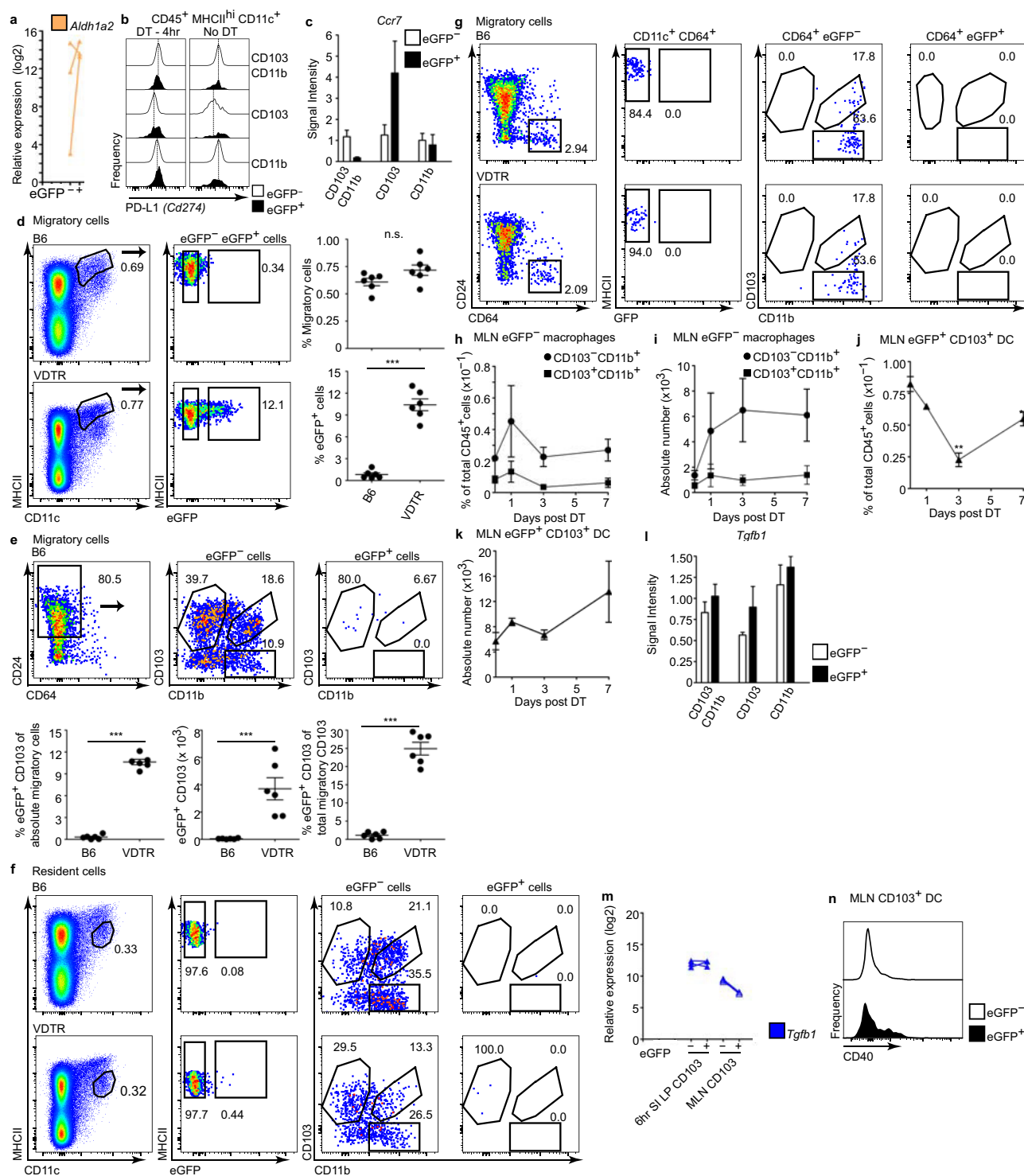
Extended Data Figure 6 | Further validation of the differential gene expression profiles following apoptotic IEC sampling. **a**, Flow cytometry validation for MER and CD205 (Ly75) protein upregulation either in the steady state or 4 h after diphtheria toxin administration, reflecting transcript data shown in Fig. 3d. Flow cytometry representative of at least three independent experiments. **b**, **g**, Hierarchical clustering of differentially expressed 'find me/eat me' receptor (**b**) and 'suppression of immune response' (**g**) genes at 4 h after diphtheria toxin administration. Differences in expression did (**g**) and did not (**b**) meet statistical significance of at least 1.2-fold (ANOVA ($Q < 0.05$) and Tukey's HSD post-hoc test ($P < 0.05$); $-1.2 > \text{fold} > 1.2$) comparing eGFP⁺ and eGFP⁻ phagocytes at 4 h following diphtheria toxin administration; expressed by row Z scale and relative expression (log₂). In **b**, upregulated genes correlating with eGFP content included *Cx3cr1* (the receptor for the 'find me' signal fractalkine) and TAM family receptor *Axl* by CD11b⁺ macrophages, as well as *Tyro3* by CD103⁺ dendritic cells. Different members of the Stabilin family of scavenger receptors, *Adgrb1*, *Stab1/2*, required for apoptotic cell clearance¹, as well as *Timd4*, which binds to exposed phosphatidylserine, were differentially expressed by macrophages and CD103⁺ dendritic cells. Data represent five independent experiments with four mice per experiment and three biological replicates. Many genes in **b** and **g** remained differentially regulated at 6 h after diphtheria toxin administration (see **d** and **e**), showing a stable profile not unique to the 4-h time point. **c**, qRT-PCR validation for *Axl* and *Timd4* in SI LP phagocytes at 4 h after diphtheria toxin administration. **d**, **e**, qRT-PCR

quantification of relative gene expression in the indicated eGFP⁻ and eGFP⁺ populations (log₂). CD11b⁺ macrophages (**d**) and CD103⁺ dendritic cells (**e**) from the SI LP at 6 h following diphtheria toxin administration. Expression of *Ccr7*, encoding the chemotactic receptor required for migration into mesenteric lymph nodes where priming of naive T cells occurs, was higher in SI LP eGFP⁺CD103⁺ dendritic cells, but not macrophages, relative to their eGFP⁻ counterparts (Extended Data Fig. 8c), and remained high in eGFP⁺CD103⁺ dendritic cells at 6 h after diphtheria toxin administration (**e**) compared to the downregulation by eGFP⁺ cells relative to eGFP⁻CD11b⁺ macrophages (**d**). Data represent two independent analytical repeats and one biological replicate consisting of three mice. **f**, Signal intensity of *Cd300a*, encoding an inhibitory receptor reported to bind apoptotic cell-exposed phosphatidylserine and suppress commensal-driven IFN β production by large intestine lamina propria CD103⁻CD11b⁺CX3CR1⁺F4/80^{lo-int} cells³, from microarray analyses conducted at 4 h after diphtheria toxin administration, showing similar levels in SI LP eGFP⁺ and eGFP⁻ macrophages and CD103⁺ dendritic cells indicated on the x axis. Data did not meet statistical significance of at least 1.2-fold (ANOVA ($Q < 0.05$) and Tukey's HSD post-hoc test ($P < 0.05$); $-1.2 > \text{fold} > 1.2$) and represent five independent experiments with 4 mice per experiment and three biological replicates. Data are mean \pm s.e.m. White and black bar graphs indicate expression from eGFP⁻ and eGFP⁺ phagocytes, respectively. **h**, **i**, Microarray validation by duplicate qRT-PCR for indicated genes represent three independent experiments.



Extended Data Figure 7 | Inflammatory cytokine and co-stimulatory genes in CD11c⁺ phagocytes are unchanged and decreased at 4 h and 6 h, respectively, following sampling of apoptotic IECs. Signal intensity of genes encoding inflammatory cytokines (a), and co-stimulatory molecules (b), from microarray analyses conducted 4 h after diphtheria toxin administration on the SILP phagocyte populations indicated on the x axes where CD103CD11b and CD11b denote CD103⁺CD11b⁺ and CD11b⁺ macrophages, and CD103 denotes CD103⁺ dendritic cells. White and black bar graphs indicate expression from eGFP⁻ and eGFP⁺

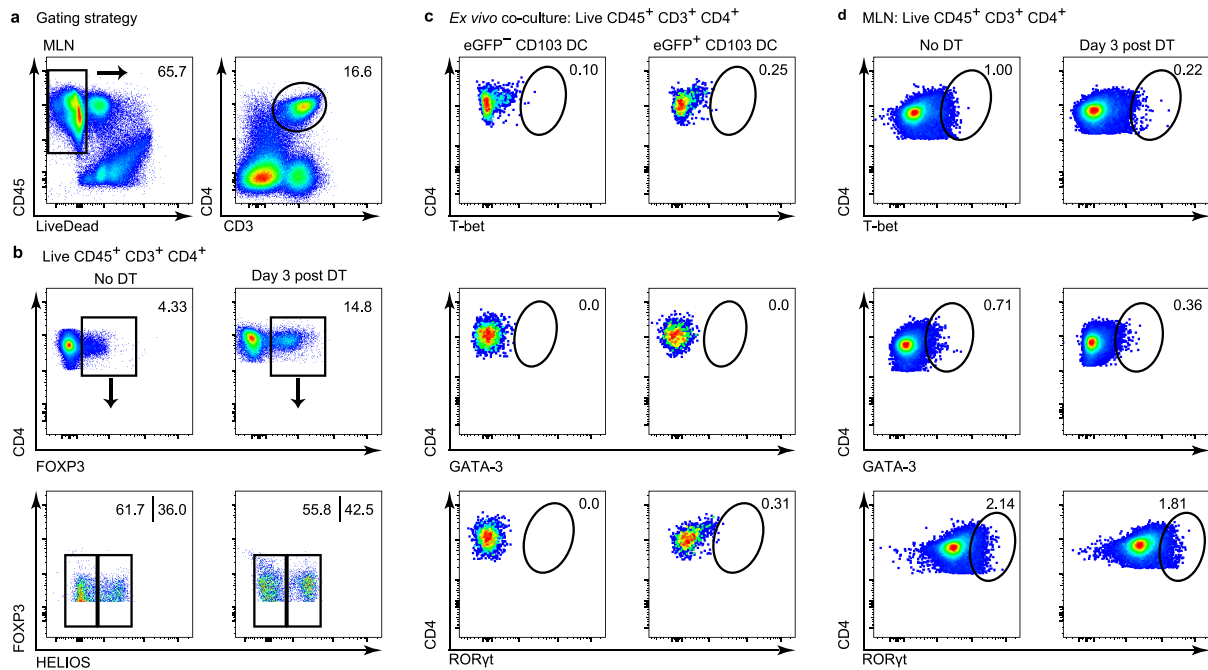
phagocytes, respectively. No statistically significant changes were observed as genes did not meet statistical significance of at least 1.2 fold (ANOVA ($Q < 0.05$) and Tukey's HSD post-hoc test ($P < 0.05$); $-1.2 > \text{fold} > 1.2$). Data represent five independent experiments with four mice per experiment and three biological replicates. **c, d**, qRT-PCR quantification of *Il1b* and *Il6* transcripts expressed by eGFP⁺ and eGFP⁻CD11b⁺ macrophages (c) and CD103⁺ dendritic cells (d) from the SILP at 6 h after diphtheria toxin administration. Data represent three independent analytical repeats and one biological replicate consisting of three mice.



Extended Data Figure 8 | See next page for caption.

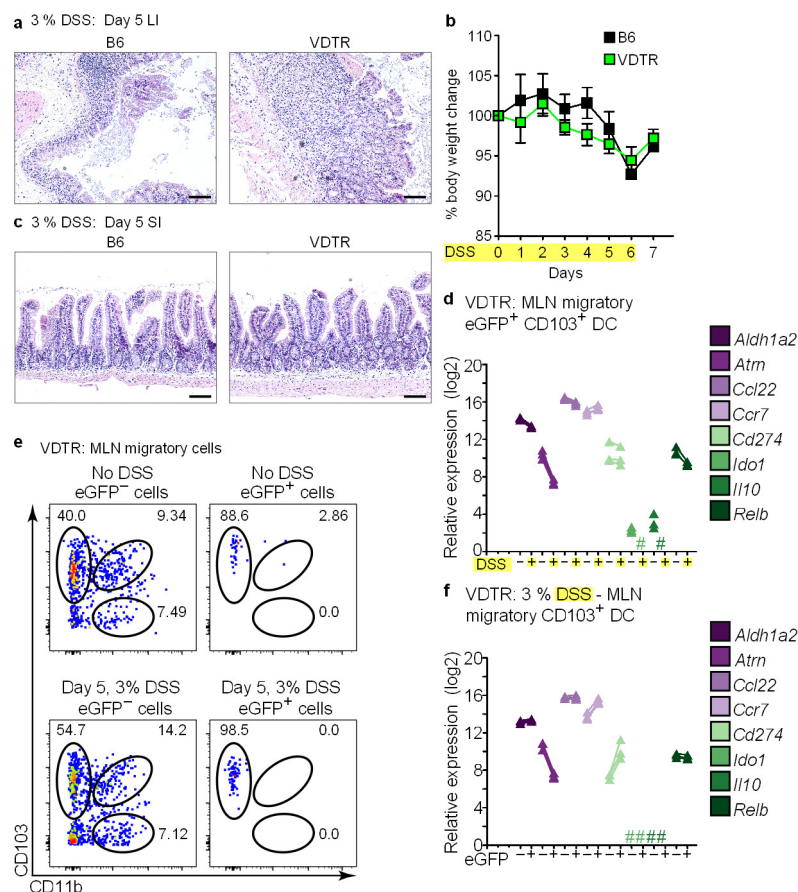
Extended Data Figure 8 | Resident phagocytes in the mesenteric lymph node do not carry eGFP⁺ apoptotic IEC cargo. **a**, Duplicate qRT-PCR quantification of *Aldh1a2* transcript levels. **b**, Flow cytometry validation for CD274 protein expression at steady state or at 4 h after diphtheria toxin administration reflecting transcript data shown in Fig. 4a. Data represent at least three independent experiments. **c**, **i**, Signal intensity of *Ccr7* (**c**) and *Tgfb1* (**i**) from microarray analyses. Expression of *Ccr7* was higher in SILP eGFP⁺CD103⁺ dendritic cells but not macrophages, relative to their eGFP[−] counterparts. Although *Ccr7* transcripts showed a higher expression trend in CD103⁺ dendritic cells, relative expression of *Ccr7* and *Tgfb1* did not meet statistical significance of at least 1.2-fold (ANOVA ($Q < 0.05$) and Tukey's HSD post-hoc test ($P < 0.05$); $-1.2 > \text{fold} > 1.2$). Data represent five independent experiments with four mice per experiment and three biological replicates. **d–g**, Flow cytometric analyses of MLN cells from naive, untreated B6 and VDTR mice. **d**, **e**, 'Migratory' cells were pre-gated on live CD45⁺MHCII^{hi}CD11c⁺ cells. Quantification in right (**d**) and bottom (**e**) panels. **f**, 'Resident' cells were pre-gated on live CD45⁺MHCII^{hi}CD11c⁺ cells and these cells were eGFP[−]. $n = 3$ mice per experiment, $n = 6$ mice per group; unpaired two-tailed t -test with Welch's correction; *** $P < 0.001$. **e**, **g**, eGFP[−] and eGFP⁺ cells from **d** were further sorted on the basis of CD64, CD24,

CD103 and CD11b expression. **e**, CD24⁺CD103⁺ dendritic cells from B6 control mice did not contain the eGFP label. **g**, The minor populations (~2%) of CD64⁺CD103⁺CD11b⁺ and CD64⁺CD11b⁺ macrophages within the migratory gate⁵ were also eGFP[−]. Data in **e**, **g**, represent three independent experiments. **h–k**, Flow cytometry analyses of migratory eGFP[−] and eGFP⁺ phagocytes in the MLN over time following a single injection of diphtheria toxin. Increasing the number of apoptotic IECs with diphtheria toxin treatment did not lead to a statistically significant increase in the frequency or absolute number of the indicated macrophage populations (**h** and **i**, respectively) or eGFP⁺CD103⁺ dendritic cells (**j** and **k**, respectively) over time within migratory MLN cells. This was expected given the non-inflammatory nature of the low dose diphtheria toxin and the small number ($\sim 0.5\text{--}2 \times 10^4$) of total CD103⁺ dendritic cell within the SILP, of which an even smaller fraction ($\sim 10\%$) are eGFP⁺ (Fig. 2d). Data represent four independent experiments with 4 mice per time point. Data mean \pm s.e.m. **m**, qRT-PCR for *Tgfb1* from indicated cell populations. Data represent three independent analytical repeats and one biological replicate consisting of 3 mice. **n**, Flow cytometry for surface CD40 expression on migratory eGFP⁺ and eGFP[−] MLN CD103⁺ dendritic cells from untreated VDTR mice representative of four independent experiments.



Extended Data Figure 9 | Naive CD4⁺ T cells co-cultured with eGFP⁺ or eGFP⁻CD103 dendritic cells from MLN do not differentiate into T-bet⁺, GATA-3⁺, or RORγt⁺ cells. **a, b, d,** Flow cytometric analyses of VDTR MLN. **a, b,** Gating strategy for total CD45⁺CD3⁺CD4⁺ T cells. **b,** T cells pre-gated as in **a** from non-treated and 3 days post-diphtheria-toxin-treated VDTR mice. Peripherally induced T_{reg} were identified as FOXP3⁺Helios⁻ cells. **c,** Flow cytometric analyses of splenic CD4⁺ T cells cultured *ex vivo* at 1×10^5 after enrichment of naive CD4⁺ T cells,

for 5 days with FACS-sorted 1×10^4 eGFP⁺ or eGFP⁻CD103⁺ MLN dendritic cells. Data represent two independent experiments. **d,** MLN CD4⁺ T cells from non-treated and day 3 post-diphtheria-toxin-treated VDTR mice. T-bet, GATA-3 and RORγt, characteristic of T_H1, T_H2 and T_H17 cells, respectively, were not present in *ex vivo* co-cultures (**c**) or in the MLN of diphtheria-toxin-treated mice (**d**). At least three independent experiments were performed for flow cytometry studies. Numbers in plots indicate the percentage of positively stained cells within each gate.



Extended Data Figure 10 | Inflammation does not alter the population of phagocytes that traffics apoptotic IEC cargo to MLN, but does alter its gene expression. B6 and VDTR mice for 5 days received drinking water supplemented with 3% DSS (yellow highlight in **b**, **d**). On day 5, DSS treatment was stopped and replaced with fresh drinking water.

a, **c**, Paraffin haematoxylin and eosin sections from the large intestine (**a**), and small intestine (**c**) of indicated mice. Scale bars, 100 μ m. **b**, Change in weight shown on the y axis monitored daily as shown on the x axis. DSS treatment of VDTR mice (green line) was accompanied by weight loss similar to that observed in B6 mice (black line), with minimal damage

to the terminal ileum, as expected. $n = 4$ mice per group. **d**, **f**, qRT-PCR for indicated transcripts expressed by migratory eGFP⁺CD103⁺ MLN dendritic cells from untreated and day 5 DSS-treated VDTR mice (**d**), and eGFP⁻ and eGFP⁺CD103⁺ MLN dendritic cell isolated from day 5 DSS treated VDTR mice (**f**). #, not detected. Data represent three independent analytical repeats and one biological replicate consisting of two mice. **e**, Flow cytometry analyses of MLN migratory cells from day-5 DSS-treated and non-DSS-treated mice. Data represent two independent experiments consisting of two mice. Numbers near gates indicate the percentage of positively stained cells.

Macrophages redirect phagocytosis by non-professional phagocytes and influence inflammation

Claudia Z. Han^{1,2}, Ignacio J. Juncadella¹, Jason M. Kinchen¹, Monica W. Buckley^{1,2}, Alexander L. Klibanov^{3,4}, Kelly Dryden⁵, Suna Onengut-Gumuscu^{6,7}, Uta Erdbrügger⁸, Stephen D. Turner⁹, Yun M. Shim³, Kenneth S. Tung^{3,10} & Kodi S. Ravichandran^{1,2}

Professional phagocytes (such as macrophages¹) and non-professional phagocytes^{2–8} (such as epithelial cells) clear billions of apoptotic cells and particles on a daily basis⁹. Although professional and non-professional macrophages reside in proximity in most tissues, whether they communicate with each other during cell clearance, and how this might affect inflammation, is not known¹⁰. Here we show that macrophages, through the release of a soluble growth factor and microvesicles, alter the type of particles engulfed by non-professional phagocytes and influence their inflammatory response. During phagocytosis of apoptotic cells or in response to inflammation-associated cytokines, macrophages released insulin-like growth factor 1 (IGF-1). The binding of IGF-1 to its receptor on non-professional phagocytes redirected their phagocytosis, such that uptake of larger apoptotic cells was reduced whereas engulfment of microvesicles was increased. IGF-1 did not alter engulfment by macrophages. Macrophages also released microvesicles, whose uptake by epithelial cells was enhanced by IGF-1 and led to decreased inflammatory responses by epithelial cells. Consistent with these observations, deletion of IGF-1 receptor in airway epithelial cells led to exacerbated lung inflammation after allergen exposure. These genetic and functional studies reveal that IGF-1- and microvesicle-dependent communication between macrophages and epithelial cells can critically influence the magnitude of tissue inflammation *in vivo*.

To address potential cross-regulation between phagocytes, we initially tested a panel of eleven soluble mediators that have been linked to tissue repair, inflammation dampening, and tissue morphogenesis¹¹, for their ability to modulate engulfment by LR73 fibroblasts, a non-professional phagocytic cell line. Of the eleven factors tested, only insulin-like growth factor 1 (IGF-1) significantly dampened apoptotic cell uptake, at concentrations reported in mouse and human serum (100–600 ng ml⁻¹)¹² (Fig. 1a, b). Other factors such as EGF, FGF2, VEGF, PDGF-AA and PDGF-BB did not alter phagocytic activity over a range of concentrations (Extended Data Fig. 1a), although each elicited early downstream signalling events (Extended Data Fig. 1b–e). The engulfment-dampening effect of IGF-1 was also seen with airway epithelial cell lines BEAS-2B and 16HBE14o-, and the endothelial cell line SVEC-40 (Fig. 1c, Extended Data Fig. 1f, g).

The IGF-1 effect was not due to masking phosphatidylserine (PtdSer) on the apoptotic cells (Fig. 1d). LR73 cells express the IGF-1 receptor (IGF-1R), and IGF-1 treatment elicited phosphorylation of Akt, a signalling molecule downstream of IGF-1R (Fig. 1b).

In blood, IGF-1 is bound to IGF-binding proteins (IGFBPs) that stabilize and sequester IGF-1 (ref. 13). Addition of IGFBP3 with IGF-1 restored the phagocytic capability of LR73 cells (Fig. 1e), suggesting that binding of active IGF-1 to IGF-1R was necessary. A neutralizing antibody against the human IGF-1R reversed the engulfment-dampening

effect of IGF-1 on BEAS-2B cells (Fig. 1f). Further, OSI-906 (a small molecule kinase inhibitor of IGF-1R currently tested in clinical trials) and the inhibitor NVP-AEW541 rescued the engulfment capacity of LR73 cells treated with IGF-1, with concomitant decrease in phosphorylation of IGF-1R and Erk (Fig. 1g, Extended Data Fig. 2a). IGF-II and insulin, which share structural similarity with IGF-1 but have lower affinities for IGF-1R¹³, could also reduce apoptotic cell uptake (Extended Data Fig. 2b, c), albeit at higher concentrations. Thus, productive signalling through the IGF-1R is necessary for IGF-1 to dampen apoptotic cell uptake.

In contrast to the inhibitory effect on apoptotic cell uptake, IGF-1 enhanced the uptake of PtdSer-containing liposomes of 150–200 nm in size (Fig. 1h). This was not seen with EGF or VEGF. IGF-1 enhancement required IGF-1R signalling, as the inhibitor OSI-906 reduced the increased liposome uptake (Fig. 1i). Thus, IGF-1 can redirect phagocytosis by non-professional phagocytes, suppressing uptake of larger apoptotic cells and enhancing internalization of smaller particles.

The IGF-1 effect was reversible, as washing phagocytes that were pre-treated with IGF-1 resulted in near-complete restoration of phagocytic activity (Fig. 1j). The effect of IGF-1 on phagocytes was also rapid, as adding IGF-1 simultaneously with apoptotic cells inhibited engulfment in a similar manner as that seen with the pre-treated cells (data not shown), which suggests IGF-1 interferes with early step(s) during phagocytosis. Blocking certain early signalling pathways downstream of IGF-1R, such as Akt, mTOR, Erk or PI3-kinase¹⁴, did not rescue the IGF-1-mediated decrease in apoptotic cell uptake (Extended Data Fig. 3a–d). Although IGF-1 can activate RhoA, which is known to decrease apoptotic cell engulfment¹⁵, inhibition of RhoA-mediated signalling did not reverse the suppression of IGF-1-mediated engulfment (Extended Data Fig. 3e, f). Overexpressing activated Rac1 (which promotes apoptotic cell engulfment^{16,17}) bypassed attenuation of apoptotic cell uptake (Fig. 1k), suggesting that IGF-1R acts at or before a step regulated by active Rac1. As apoptotic cell uptake requires both Rac1-dependent actin polymerization and de-polymerization¹⁸, we explored actin regulation via IGF-1. Cytochalasin D (cytoD), which promotes actin depolymerization, potentially inhibited phagocytosis of apoptotic cells³ (data not shown). Although cytoD did not affect the basal uptake of liposomes by LR73 cells, it blocked the IGF-1-induced increase in liposome uptake (Fig. 1l). Notably, cytoD-treated LR73 cells appeared morphologically normal for the duration of the assay. Latrunculin A, which promotes actin depolymerization through a different mechanism, also reversed IGF-1-mediated enhancement of liposome uptake, without affecting the basal uptake of liposomes (Fig. 1m). Additionally, Arp2/3 complex can regulate phagocytosis through the formation of branched actin networks; however, CK-666, a small molecule inhibitor of Arp2/3, had minimal effect on the IGF-1-mediated increase of liposome uptake (Extended Data Fig. 3g).

¹The Center for Cell Clearance, University of Virginia, Charlottesville, Virginia 22903, USA. ²Department of Microbiology, Immunology, and Cancer Biology, University of Virginia, Charlottesville, Virginia 22903, USA. ³Department of Medicine, University of Virginia, Charlottesville, Virginia 22903, USA. ⁴Robert M. Berne Cardiovascular Research Center, University of Virginia, Charlottesville, Virginia 22903, USA. ⁵Department of Molecular Physiology and Biological Physics, University of Virginia, Charlottesville, Virginia 22903, USA. ⁶Department of Medicine, Division of Endocrinology and Metabolism, University of Virginia, Charlottesville, Virginia 22903, USA. ⁷Center for Public Health Genomics, University of Virginia, Charlottesville, Virginia 22903, USA. ⁸Department of Medicine, Division of Nephrology, University of Virginia, Charlottesville, Virginia 22903, USA. ⁹Department of Public Health Sciences, University of Virginia, Charlottesville, Virginia 22903, USA. ¹⁰Department of Pathology, University of Virginia, Charlottesville, Virginia 22903, USA.

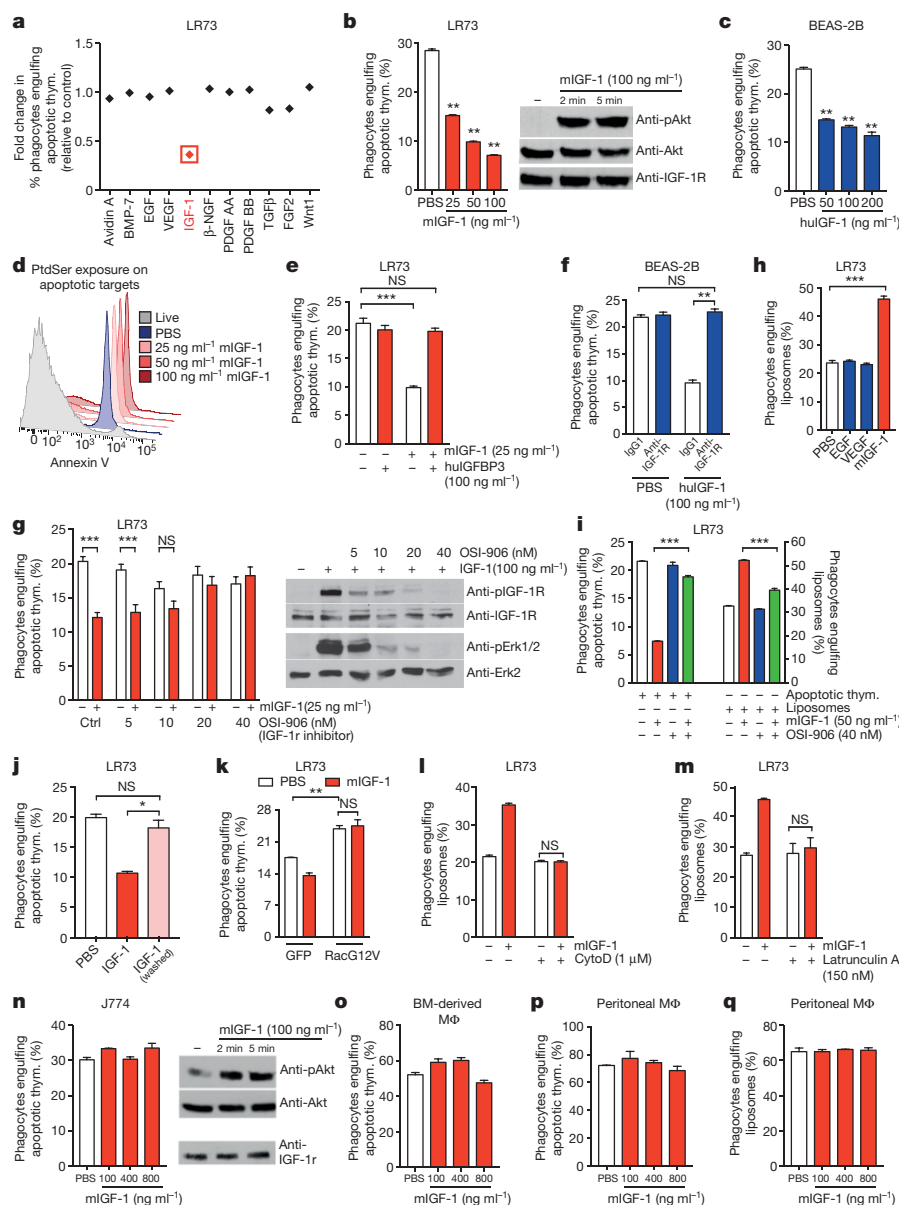


Figure 1 | IGF-1 dampens apoptotic cell engulfment and enhances liposome uptake by non-professional phagocytes. **a**, Effect of 11 growth factors on apoptotic cell engulfment by LR73 cells (each normalized to vehicle control). Thym., thymocytes. **b**, Left, engulfment by LR73 cells treated with increasing concentrations of IGF-1. Right, immunoblotting for IGF-1R and Akt phosphorylation. **c**, Apoptotic cell uptake by BEAS-2B cells treated with human IGF-1. **d**, IGF-1 does not affect Annexin V binding to apoptotic thymocytes ($n = 6$). **e**, **f**, Reversal of IGF-1-mediated engulfment inhibition of LR73 cells by IGF-1R-neutralizing antibody (**f**). **g**, IGF-1-mediated reduction in engulfment is reversed by the IGF-1R inhibitor OSI-906 (left), and immunoblotting for IGF-1R and Erk1/2 phosphorylation (right). **h**, Liposome uptake by LR73 cells treated with either IGF-1 (50 ng ml⁻¹), EGF or VEGF (100 ng ml⁻¹ each). **i**, Enhanced liposome uptake by IGF-1 is reversed by OSI-906. **j**, LR73 cells pretreated with IGF-1 were washed, and incubated with apoptotic thymocytes in the presence or absence of IGF-1. **k**, Engulfment by LR73 cells transfected with RacG12V treated and with mIGF-1. **l**, **m**, Engulfment by LR73 cells treated with cytoD (**l**) or latrunculin A (**m**) and incubated with liposomes with or without IGF-1. **n**, Left, phagocytosis of apoptotic thymocytes by J774 macrophage cells treated with IGF-1. Right, immunoblotting for IGF-1R expression and Akt phosphorylation. **o**, **p**, Apoptotic cell engulfment by bone-marrow-derived macrophages (M ϕ) or resident peritoneal macrophages treated with IGF-1. **q**, Liposome uptake by resident peritoneal macrophages treated with IGF-1. For **a–q**, except **d**, $n = 3$; representative experiment is shown and data are mean \pm s.d. NS, not significant. P value of <0.05 (*), <0.01 (**), or <0.001 (***). See Supplementary Fig. 1 for uncropped immunoblots.

Collectively, IGF-1-mediated modulation of phagocytosis involves rapid and reversible modification of F-actin/G-actin dynamics, but probably not Arp2/3-mediated functions.

We next tested the effect of IGF-1 on professional phagocytes. Adding IGF-1 to the macrophage cell lines J774 or IC-21, or primary macrophages (bone-marrow-derived or resident peritoneal macrophages), did not affect apoptotic cell uptake, even with supra-physiological concentrations of IGF-1, at every time point analysed (Fig. 1n–p, Extended Data Fig. 4a, d). Peritoneal macrophages exposed to IGF-1 also had comparable liposome uptake as control-treated cells (Fig. 1q). Importantly, IGF-1R expression was confirmed in all macrophage cell types analysed (Extended Data Fig. 4c) and IGF-1 initiated early signalling in macrophages, as assessed by IGF-1R and Akt phosphorylation (Fig. 1n, Extended Data Fig. 4a). These data suggest that IGF-1-mediated modulation of phagocytosis does not extend to macrophages.

High levels of serum IGF-1 is linked to IGF-1 production in the liver, but macrophages can also produce IGF-1 in response to the cytokine IL-4 (ref. 19), raising the possibility that within a tissue, IGF-1 release from macrophages could regulate nearby non-professional phagocytes. We first confirmed IGF-1 secretion by peritoneal macrophages treated with IL-4 (Fig. 2a). Furthermore, resident peritoneal macrophages exposed to apoptotic Jurkat cells (but not live cells) produced IGF-1

(Fig. 2a). IGF-1 protein induction appears to be from newly transcribed *Igf1* mRNA (Extended Data Fig. 5a). Macrophage-produced IGF-1 could also suppress apoptotic cell uptake by LR73 cells, and this was reversed by addition of IGFBP3 (Fig. 2b).

We chose to examine lung tissue to investigate *in vivo* communication between phagocytes, as alveolar macrophages and airway epithelial cells reside in close proximity, and both engulf apoptotic cells³. Alveolar macrophages, similar to resident peritoneal macrophages, are derived from fetal monocytes²⁰, and are readily isolated through CD11c⁺Siglec F⁺ expression, whereas the airway epithelial cells can be isolated and tracked with available genetic tools³. IGF-1R expression was prominent in the airway epithelial cells (Extended Data Fig. 5b, c), but was also detectable at lower levels in alveolar macrophages. To test the effect of IGF-1 *in vivo*, we used mice in which airway epithelial cells are specifically marked by yellow fluorescence protein (YFP). When apoptotic cells or liposomes were administered intranasally with IGF-1 (Fig. 2c), YFP⁺ airway epithelial cells had decreased apoptotic cell engulfment, but enhanced liposome uptake (Fig. 2d, e). Importantly, lung macrophages from the same IGF-1-treated mice were unaffected in their ability to engulf these targets (Fig. 2d, e). Thus, IGF-1 can differentially affect engulfment by professional and non-professional phagocytes that reside nearby in the same tissue.

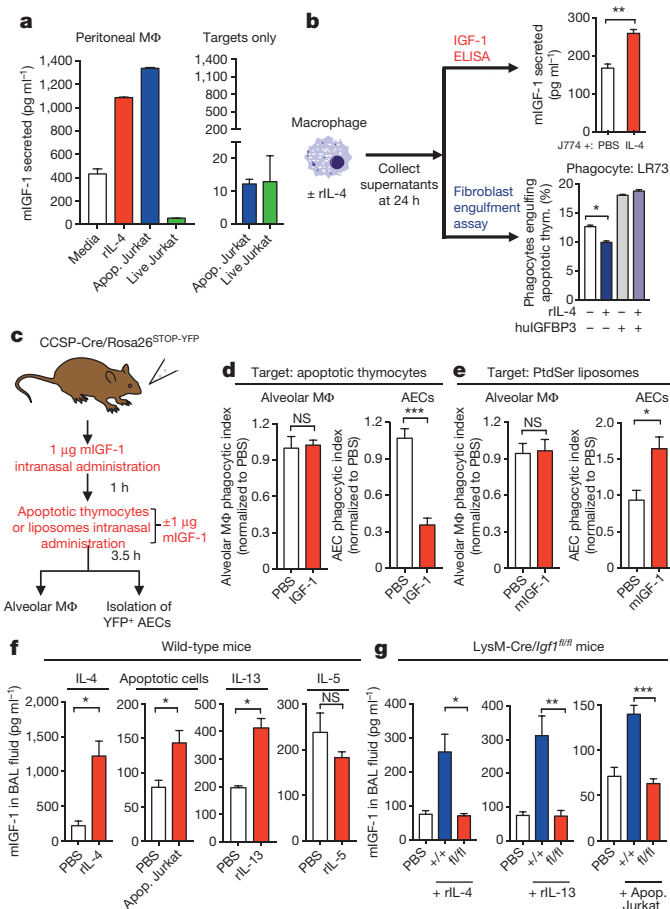


Figure 2 | Macrophages produce IGF-1 during apoptotic cell clearance.

a, Left, IGF-1 secretion by peritoneal macrophages stimulated with IL-4, apoptotic Jurkat cells or live Jurkat cells. Right, apoptotic or live Jurkat cells do not produce IGF-1 (representative of $n = 3$), mean \pm s.d. **b**, J774 cells were treated with recombinant IL-4 (rIL-4). After 24 h, half of the supernatant was assessed for IGF-1 secretion by ELISA, (right panel) and the other half was incubated with PBS or IGFBP3 and then tested in the phagocytosis assay using LR73 cells (right) (representative of $n = 3$), mean \pm s.d. Note, addition of IGFBP3 increases the basal phagocytosis owing to low basal levels of IGF-1 found in the supernatant of unstimulated J774 cells. **c–e**, CCSP-Cre mice with YFP⁺ Club cells were pre-treated with 1 μg IGF-1 intranasally for 1 h and then administered targets with or without IGF-1. Uptake of apoptotic thymocytes (**d**, $n = 6$ per group) or liposomes (**e**, $n = 3$ per group) by alveolar macrophages and airway epithelial cells were then assessed. **f**, Wild-type mice were given PBS, IL-4, apoptotic cells, IL-5, or IL-13 intranasally for 2 consecutive days, and the BAL fluid assessed for IGF-1 ($n = 2–3$ mice per group for cytokines, $n = 5, 8$ for apoptotic cell instillation). **g**, LysM-Cre/Igf1^{fl/fl} or Igf1^{fl/fl} mice were given PBS, IL-4 or IL-13, or apoptotic cells intranasally, and BAL fluid assessed for IGF-1 ($n = 6, 6, 4$ mice per group for rIL-4; $n = 6, 4, 4$ for rIL-13; $n = 6, 9, 9$ mice per group for apoptotic cell instillation). Data are mean \pm s.e.m. unless otherwise indicated. NS, not significant. P value of <0.05 (*), <0.01 (**), or <0.001 (***).

We next examined IGF-1 release by alveolar macrophages *in vivo*. Intranasal administration of IL-4 or apoptotic Jurkat cells resulted in an increased level of IGF-1 in the bronchoalveolar lavage (BAL) fluid (Fig. 2f). When we tested IL-13 or IL-5 (cytokines linked to lung homeostasis and inflammation²¹), IL-13 (whose receptor shares a common subunit with the IL-4R), but not IL-5, induced IGF-1 production (Fig. 2f). As the source of IGF-1 is difficult to distinguish in these experiments, we generated LysM-Cre/Igf1^{fl/fl} mice, which express Cre under the control of the endogenous *LysM* (also known as *Ly22*) promoter, and target IGF-1 deletion in the myeloid lineage. LysM-Cre/Igf1^{fl/fl} mice showed loss of *Igf1* mRNA in alveolar macrophages

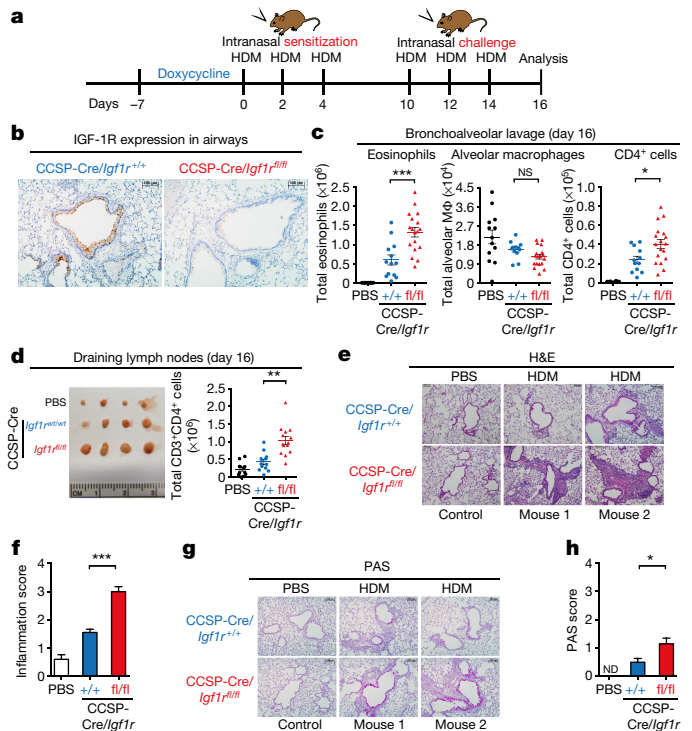


Figure 3 | Mice lacking IGF-1R in airway epithelial cells have exacerbated airway inflammation.

a, Schematic of HDM-induced allergic airway inflammation. **b**, Representative images showing IGF-1R expression in bronchial epithelial cells, and its loss in CCSP-rtTA/tetO-Cre/Igf1^{fl/fl} mice treated with doxycycline. **c**, Numbers of eosinophils, alveolar macrophages, and CD4⁺ T cells in the BAL fluid of CCSP-Cre/Igf1^{fl/fl} and Igf1^{fl/fl} mice administered PBS or HDM (each dot represents a mouse). **d**, Left, representative lung-draining lymph nodes from CCSP-Cre/Igf1^{fl/fl} and CCSP-Cre/Igf1^{fl/fl} mice that were given PBS or HDM. Right, total CD4⁺ T-cell counts from lymph nodes. **e–h**, Representative haematoxylin and eosin (H&E) images (**e**) or periodic acid–Schiff (PAS) staining (**g**) of lung sections from CCSP-Cre/Igf1^{fl/fl} and CCSP-Cre/Igf1^{fl/fl} mice given PBS or HDM ($n = 3–4$ mice per condition). Representative histological scoring of inflammation (**f**) and PAS staining (**h**) ($n = 6–10$ sections and 3 mice per condition). All data are presented as mean \pm s.e.m. NS, not significant. P value of <0.05 (*), <0.01 (**), or <0.001 (***).

(Extended Data Fig. 5d) and IGF-1 induction after treatment with IL-4, IL-13 or apoptotic cells (Fig. 2g). This suggested that the macrophage/myeloid population is the predominant source of inducible IGF-1 in the lung.

To investigate IGF-1–IGF-1R signalling in inflammation, we used a model of airway inflammation induced by the allergen house dust mite (HDM)²¹, where apoptotic cell recognition by airway epithelial cells influences inflammation³. We administered HDM intranasally in the allergen sensitization and challenge stage, modelling the natural route of allergen encounter (Fig. 3a). When we first tested LysM-Cre/Igf1^{fl/fl} mice, these mice proved unsuitable, as we continued to detect variable levels of IGF-1 in the BAL fluid after HDM administration (probably owing to leakage of serum IGF-1 during inflammation). Therefore, we targeted instead IGF-1R on the epithelial cells. We crossed Igf1^{fl/fl} mice with CCSP-rtTA/tetO-Cre transgenic mice (see Methods), the latter driving Cre under the Club cell secretory protein (CCSP; also known as *Scgblal*) promoter in the epithelial cells of the trachea, bronchi and bronchioles. This CCSP-rtTA/tetO-Cre strain allows for inducible Cre expression through doxycycline administration through drinking water, thereby allowing normal development and gene deletion just before allergen exposure. We observed near-complete loss of IGF-1R on epithelial cells of CCSP-Cre/Igf1^{fl/fl} mice after doxycycline treatment (Fig. 3b). After sensitization and challenge with low-endotoxin HDM,

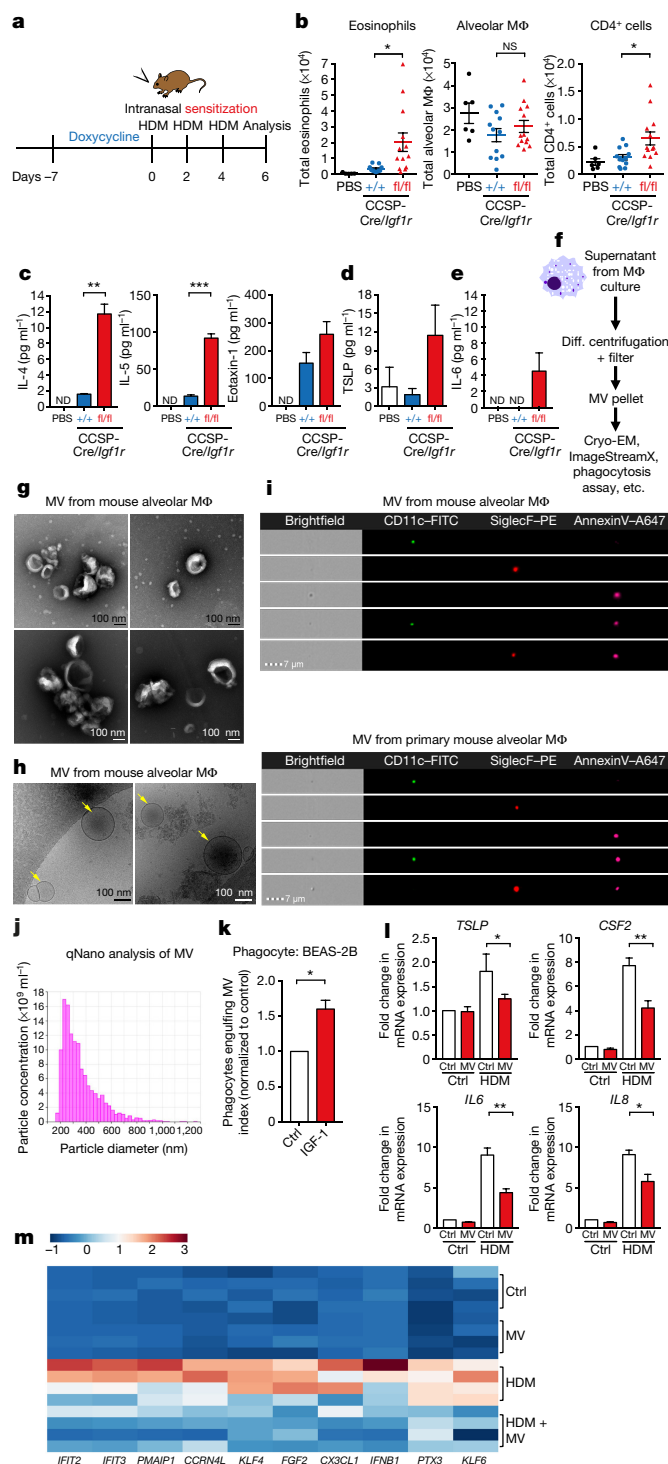


Figure 4 | IGF-1R expression in airway epithelial cells is required in the sensitization phase of airway inflammation. **a**, Schematic of IGF-1R deletion before the sensitization phase to assess its effect on early stages of inflammation. **b**, Numbers of eosinophils, alveolar macrophages, and CD4⁺ T cells on day 6 in the BAL of CCSP-Cre/*Igf1r*^{+/+} and CCSP-Cre/*Igf1r*^{fl/fl} mice primed with PBS or HDM. **c–e**, Analysis of IL-4, IL-5, eotaxin-1, and IL-6 (via Luminex (**c**, **e**), *n* = 3 mice per group) and TSLP (by ELISA (**d**), *n* = 2, 7, 9 mice per group) in the BAL fluid from representative CCSP-Cre/*Igf1r*^{+/+} and CCSP-Cre/*Igf1r*^{fl/fl} mice primed with PBS or HDM. **f**, Schematic of generation and isolation of alveolar-macrophage-derived microvesicles. **g, h**, Representative negative-stain EM (**g**) or cryo-EM (**h**) images of microvesicles isolated from mouse alveolar macrophages. Images show spherical membrane-bound structures of a range of sizes (yellow arrows). **i**, Representative ImageStreamX images of microvesicles isolated from mouse alveolar macrophage cell line and primary mouse alveolar macrophages and stained for alveolar macrophage markers. **j**, Tunable resistive pulse sensing analysis of microvesicles from alveolar macrophages using qNano, pore size 400 nm, to determine frequency and sizing of microvesicles (representative of *n* = 3). **k**, BEAS-2B cells treated with IGF-1 (100 ng ml⁻¹) were assessed for uptake of alveolar-macrophage-derived microvesicles (MV) (*n* = 4). **l**, BEAS-2B cells were treated with HDM either in the presence or absence of alveolar-macrophage-derived microvesicles (MV) for 3 h and assessed for expression of TSLP, CSF2, IL6, IL8 (*n* = 4). **m**, Heat map of 10 most differentially expressed genes from RNA-seq analysis of BEAS-2B cells exposed to HDM with or without alveolar macrophage-derived microvesicles. Data presented as mean ± s.e.m. ND, not detected, NS, not significant. P value of <0.05 (*), <0.01 (**), or <0.001 (***).

data suggested a requirement for IGF-1R in the airway epithelial cells in minimizing inflammation in this model.

These observations were initially surprising, as we expected the loss of IGF-1R on airway epithelial cells to improve apoptotic cell clearance and thereby attenuate, rather than worsen, inflammation. This prompted us to examine the temporal requirement of IGF-1–IGF-1R signalling in epithelial cells during allergen exposure. To distinguish the requirement for IGF-1–IGF-1R signalling at the sensitization versus challenge phase, we administered doxycycline at different times: deleting IGF-1R expression before allergen sensitization (Fig. 4a), or after the initial allergen sensitization but before the challenge phase (Extended Data Fig. 7a). Mice with IGF-1R deletion before allergen sensitization showed a significant increase in inflammatory parameters (Fig. 4b), whereas deletion of IGF-1R after the sensitization phase had minimal effect on disease severity (Extended Data Fig. 7b, c).

Airway epithelial cells that encounter allergens can produce cytokines, such as TSLP and CSF-2/GM-CSF (that affect dendritic cell maturation), as well as IL-33 and IL-25 that drive type-2 innate lymphoid cells (ILC2s) to proliferate and produce IL-13 (ref. 21). Human bronchial epithelial cells from individuals with asthma produce elevated levels of IL-6, IL-8, and CSF-2 (ref. 22), and TSLP released from airway epithelial cells (mice and human) can exacerbate airway inflammation²³. When we assessed the cytokine profile in the BAL fluid of CCSP-Cre/*Igf1r*^{fl/fl} mice sensitized with HDM, levels of TSLP and IL-6 were greater (Fig. 4d, e). We did not find increased levels of IL-33, as seen with Rac1 deletion in airway epithelial cells³, but this may be due to differences in inflammatory pathways regulated by the loss of IGF-1R rather than Rac1. These results suggest that IGF-1R in airway epithelial cells limits inflammatory cytokines during the initial antigen exposure/sensitization phase.

Regarding trigger(s) for IGF-1 release from the alveolar macrophages in the sensitization phase, although IL-4 is often considered a classic T_H2 cytokine from the adaptive immune response²¹, both IL-4 (refs 24, 25) and IL-13 (ref. 21) can be produced during the allergen sensitization phase by resident mast cells, basophils, and/or ILC2s. In fact, IL-4 was enhanced (along with IL-5 and eotaxin-1) in the BAL fluid of CCSP-Cre/*Igf1r*^{fl/fl} mice sensitized with HDM (Fig. 4c). Thus, IL-4 and/or apoptotic cells present during the initial allergen exposure could elicit IGF-1 production from alveolar macrophages.

the CCSP-Cre/*Igf1r*^{fl/fl} mice had greater airway inflammation based on several parameters. (1) There was marked increase in eosinophils and CD4⁺ T cells in the BAL fluid (Fig. 3c) and trending increase in inflammatory cells in the lungs (Extended Data Fig. 6a). (2) Lung-draining lymph nodes were larger and contained more CD4⁺ T cells (Fig. 3d). (3) Lung sections showed increased peribronchial and perivascular cellular infiltration (Fig. 3e, f) and greater mucus accumulation after HDM treatment (Fig. 3g, h). (4) HDM-treated CCSP-Cre/*Igf1r*^{fl/fl} mice displayed increased airway reactivity after methacholine challenge, a measure of the bronchial hyper-responsiveness (Extended Data Fig. 6b). There were more apoptotic cells in lung sections (cleaved caspase 3 staining) in HDM-treated CCSP-Cre/*Igf1r*^{fl/fl} mice, probably owing to greater inflammation (Extended Data Fig. 6c). These

Alveolar macrophages are reported to release microvesicles containing anti-inflammatory mediators during smoking-induced lung injury²⁶. As IGF-1 enhances liposome uptake by airway epithelial cells, we tested whether microvesicles from alveolar macrophages may affect airway epithelial cell response to HDM. Using differential centrifugation and filtration, we isolated and characterized microvesicles from alveolar macrophages (Fig. 4f). Negative stain and cryo-electron microscopy revealed membrane-bound spherical structures >100 nm in diameter (Fig. 4g, h). Through ImageStreamX analysis, we determined that the particles from an alveolar macrophage cell line or primary mouse alveolar macrophages carried markers of lung macrophage origin, CD11c or Siglec F, with variable Annexin V staining (Fig. 4i). Tunable resistive pulse sensing analysis of microvesicles revealed a mean particle size of 357 ± 148.5 nm (Fig. 4j), within the reported 100–1,000 nm range for microvesicles²⁷. Further, IL-4 treatment of macrophages increased microvesicles secretion (Extended Data Fig. 8a). In addition, uptake of these microvesicles by BEAS-2B airway epithelial cells was enhanced by IGF-1 (Fig. 4k). Notably, adding alveolar macrophage-derived microvesicles to HDM-treated BEAS-2B cells resulted in significantly lower *TSLP*, *CSF2*, *IL6*, and *IL8* induction (Fig. 4l). To further explore this, we performed RNA-seq of BEAS-2B cells treated with HDM \pm microvesicles. Whereas HDM treatment increased several known genes associated with asthma in humans^{28–30}, such as *FGF2*, *KLF4*, *IFIT2*, and *PTX3*, adding microvesicles from alveolar macrophages suppressed transcription of these genes in the epithelial cells (Fig. 4m, Extended Data Fig. 8b). Thus, microvesicles released from alveolar macrophages acting on airway epithelial cells exposed to allergen could dampen airway inflammation.

The data presented here provide several insights into phagocytosis and tissue inflammation. It has long been known that professional and non-professional phagocytes reside in proximity and can engulf dying cells and debris, yet communication between these phagocytes was not understood. The data presented here identifies a rapid, transient, and reversible regulation, wherein soluble IGF-1 from macrophages influences the type of particle taken up by epithelial cells. This transient effect might allow the macrophages to temporarily ‘redirect’ the non-professional phagocytes towards other function(s). Such a possible hierarchy in cell clearance could provide temporal and spatial cross-communication within a given tissue⁹. Our data identifies a two-part regulation of epithelial cells by macrophages that affects airway inflammation, that is, the secretion of IGF-1 redirects particle uptake, and the release of microvesicles dampens inflammatory cytokine production by epithelial cells (Extended Data Fig. 9). IGF-1 is a growth factor widely linked to growth, cellular proliferation, and ageing. Global IGF-1 deletion in mice results in dwarfism and perinatal lethality, and IGF-1 mutations are linked to human diseases¹³. This work identifies a previously unappreciated biological function for the IGF-1–IGF-1R signalling axis as a modulator of airway hyper-responsiveness to allergens with potential therapeutic relevance.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 25 July; accepted 4 October 2016.

Published online 7 November 2016.

- Henson, P. M. & Hume, D. A. Apoptotic cell removal in development and tissue homeostasis. *Trends Immunol.* **27**, 244–250 (2006).
- Burstyn-Cohen, T. *et al.* Genetic dissection of TAM receptor-ligand interaction in retinal pigment epithelial cell phagocytosis. *Neuron* **76**, 1123–1132 (2012).
- Juncadella, I. J. *et al.* Apoptotic cell clearance by bronchial epithelial cells critically influences airway inflammation. *Nature* **493**, 547–551 (2013).
- Wood, W. *et al.* Mesenchymal cells engulf and clear apoptotic footplate cells in macrophageless PU.1 null mouse embryos. *Development* **127**, 5245–5252 (2000).
- Lu, Z. *et al.* Phagocytic activity of neuronal progenitors regulates adult neurogenesis. *Nat. Cell Biol.* **13**, 1076–1083 (2011).
- Ichimura, T. *et al.* Kidney injury molecule-1 is a phosphatidylserine receptor that confers a phagocytic phenotype on epithelial cells. *J. Clin. Invest.* **118**, 1657–1668 (2008).

- Mesa, K. R. *et al.* Niche-induced cell death and epithelial phagocytosis regulate hair follicle stem cell pool. *Nature* **522**, 94–97 (2015).
- Elliott, M. R. *et al.* Unexpected requirement for ELMO1 in clearance of apoptotic germ cells *in vivo*. *Nature* **467**, 333–337 (2010).
- Gregory, C. D. & Pound, J. D. Microenvironmental influences of apoptosis *in vivo* and *in vitro*. *Apoptosis* **15**, 1029–1049 (2010).
- Green, D. R. The end and after: how dying cells impact the living organism. *Immunity* **35**, 441–444 (2011).
- Werner, S. & Grose, R. Regulation of wound healing by growth factors and cytokines. *Physiol. Rev.* **83**, 835–870 (2003).
- Yakar, S. *et al.* Circulating levels of IGF-1 directly regulate bone growth and density. *J. Clin. Invest.* **110**, 771–781 (2002).
- Smith, T. J. Insulin-like growth factor-I regulation of immune function: a potential therapeutic target in autoimmune diseases? *Pharmacol. Rev.* **62**, 199–236 (2010).
- LeRoith, D. Insulin-like growth factor I receptor signaling—overlapping or redundant pathways? *Endocrinology* **141**, 1287–1288 (2000).
- Tosello-Trampont, A. C., Nakada-Tsukui, K. & Ravichandran, K. S. Engulfment of apoptotic cells is negatively regulated by Rho-mediated signaling. *J. Biol. Chem.* **278**, 49911–49919 (2003).
- Leverrier, Y. & Ridley, A. J. Requirement for Rho GTPases and PI 3-kinases during apoptotic cell phagocytosis by macrophages. *Curr. Biol.* **11**, 195–199 (2001).
- Tosello-Trampont, A. C., Brugnera, E. & Ravichandran, K. S. Evidence for a conserved role for CRKII and Rac in engulfment of apoptotic cells. *J. Biol. Chem.* **276**, 13797–13802 (2001).
- Schlam, D. *et al.* Phosphoinositide 3-kinase enables phagocytosis of large particles by terminating actin assembly through Rac/Cdc42 GTPase-activating proteins. *Nat. Commun.* **6**, 8623 (2015).
- Chen, F. *et al.* An essential role for T_H2-type responses in limiting acute tissue damage during experimental helminth infection. *Nat. Med.* **18**, 260–266 (2012).
- Ginhoux, F. & Williams, M. Tissue-resident macrophage ontogeny and homeostasis. *Immunity* **44**, 439–449 (2016).
- Lambrecht, B. N. & Hammad, H. The immunology of asthma. *Nat. Immunol.* **16**, 45–56 (2015).
- Marinari, M., Vittori, E., Hollemborg, J. & Mattoli, S. Expression of the potent inflammatory cytokines, granulocyte-macrophage-colony-stimulating factor and interleukin-6 and interleukin-8, in bronchial epithelial cells of patients with asthma. *J. Allergy Clin. Immunol.* **89**, 1001–1009 (1992).
- Zhou, B. *et al.* Thymic stromal lymphopoietin as a key initiator of allergic airway inflammation in mice. *Nat. Immunol.* **6**, 1047–1053 (2005).
- Gessner, A., Mohrs, K. & Mohrs, M. Mast cells, basophils, and eosinophils acquire constitutive IL-4 and IL-13 transcripts during lineage differentiation that are sufficient for rapid cytokine production. *J. Immunol.* **174**, 1063–1072 (2005).
- Sokol, C. L., Barton, G. M., Farr, A. G. & Medzhitov, R. A mechanism for the initiation of allergen-induced T helper type 2 responses. *Nat. Immunol.* **9**, 310–318 (2008).
- Bourdonnay, E. *et al.* Transcellular delivery of vesicular SOCS proteins from macrophages to epithelial cells blunts inflammatory signaling. *J. Exp. Med.* **212**, 729–742 (2015).
- Burger, D. *et al.* Microparticles: biomarkers and beyond. *Clin. Sci. (Lond.)* **124**, 423–441 (2013).
- Bissonnette, E. Y. *et al.* Fibroblast growth factor-2 is a sputum remodeling biomarker of severe asthma. *J. Asthma* **51**, 119–126 (2014).
- Zhang, J. *et al.* Pentraxin 3 (PTX3) expression in allergic asthmatic airways: role in airway smooth muscle migration and chemokine production. *PLoS One* **7**, e34965 (2012).
- Chamberland, A., Madore, A. M., Tremblay, K., Laviolette, M. & Laprise, C. A comparison of two sets of microarray experiments to define allergic asthma expression pattern. *Exp. Lung Res.* **35**, 399–410 (2009).

Supplementary Information is available in the online version of the paper.

Acknowledgements The authors thank members of the Ravichandran laboratory for discussions and critical reading of manuscript; S. Arandjelovic, K. Penberthy, and L. Haney for scoring histology slides, and J. Perry for assistance with generation of the heat map, and the UVA Flow Cytometry Core and the UVA Molecular Electron Microscopy Core. This work is supported by grants to K.S.R. from the NIGMS (GM064709), NIMH (MH096484), NHLBI (P01HL120840), and Center for Cell Signalling at the University of Virginia, NHLBI (HL132287 and HL091127) to Y.M.S. and K23 HL12610 to U.E. Additional support was provided via the NIH Training Grants T32 GM008136 (Cell and Molecular Biology) to C.Z.H.; T32 AI007496 (Immunology) to C.Z.H., M.W.B., and I.J.J.; and T32 GM007267 (MSTP) to M.W.B.

Author Contributions C.Z.H. designed, performed and analysed most of the experiments in this study with input from K.S.R. I.J.J., J.M.K., M.W.B. assisted with *in vivo* experiments. A.L.K. provided phosphatidylserine liposomes. K.D. performed EM imaging of microvesicles. S.O.G. ran the RNA-seq. U.E. performed microvesicle quantification using qNano. S.D.T. analysed the RNA-seq data. Y.M.S. performed airway resistance experiments. K.S.T. assisted in evaluation of lung pathology. C.Z.H. and K.S.R. wrote the manuscript with input from co-authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to K.S.R. (ravi@virginia.edu).

Reviewer Information *Nature* thanks L. Hedrick and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

Mice. C57BL/6J, *Igf1^{fl/fl}*, and *Igf1^{fl/jf}* mice were obtained from Jackson Laboratories. CCSP-rtTA/tetO-Cre were provided by J. Whitsett at Cincinnati Children's Hospital³¹. To generate the IGF-1R deletion in Club cells, we crossed CCSP-rtTA/tetO-Cre mice to *Igf1^{fl/jf}* (ref. 32). To achieve deletion, mice were given doxycycline (1 mg ml⁻¹) in drinking water containing 0.4% sucrose for at least 7 days before beginning of allergen administration, unless otherwise noted. Administering doxycycline, a derivative of tetracycline, binds to reverse tetracycline-controlled transactivator (rtTA), which then allows rtTA to bind the tetracycline response element, resulting in Cre expression. We also crossed *Igf1^{fl/jf}* mice^{33,34} to LysM-Cre mice to conditionally delete *Igf1* in the myeloid lineage. We have previously reported on the generation of CCSP-Cre/YFP⁺ mice³. For all *in vivo* experiments, except generation of apoptotic thymocytes, mice between the ages of 8 and 12 weeks were used. Mice were allocated to experimental groups on the basis of genotype and age-matching. Female and male mice were used for all experiments, except for *in vivo* engulfment assays in which only male mice were used. All animal procedures were performed according to the protocols provided by the Institutional Animal Care and Use Committee (IACUC) of the University of Virginia. No statistical tests were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Induction of airway inflammation. Mice were given drinking water containing doxycycline (1 mg ml⁻¹) seven days before first HDM administration. Mice were primed intranasally with 10 µg of low endotoxin house dust mite extraction (Indoor Biotechnologies) on days 0, 2, 4 and then challenged intranasally on days 10, 12, and 14. On day 16, mice were collected and analysed for eosinophilic airway inflammation. Alternatively, mice were given three doses of low endotoxin HDM on days 0, 2, and 4 and analysed on day 6 ('sensitization phase'). For the 'challenge phase', mice that had not received any doxycycline were given three doses of low endotoxin HDM on days 0, 2 and 4, then given doxycycline (in drinking water) from day 4 until the mice were analysed. These mice were also challenged intranasally with three doses of low endotoxin HDM on days 10, 12, and 14 and analysed on day 16.

Collection of BAL fluid, lymph nodes, and lung. For airway inflammation experiments, 0.8 ml of PBS were delivered intratracheally through a cannula. Recovered BAL fluid was centrifuged and the supernatant was frozen at -80 °C for subsequent Luminex analysis. Collected cells were stained for surface markers to distinguish macrophages, neutrophils, T cells, and eosinophils. For lung collection, mice were perfused through the right ventricle with PBS and the lungs were carefully excised and placed in type-2 collagenase (Worthington Biochemical Corporation) dissolved in HBSS containing Ca²⁺ and Mg²⁺. Lungs were minced and then incubated at 37 °C for one h, with vigorous pipetting every 15 min to separate the tissue. The lung homogenate was then passed through a 70-µm nylon strainer, spun down and treated with red blood cell lysis buffer (Sigma-Aldrich) for 5 min. The cells were then washed and resuspended in PBS containing 0.1% BSA. Draining lymph nodes were carefully extracted, and a single cell suspension was made by passage through a 70-µm nylon strainer using the flat end of a syringe. Cells were washed and then resuspended in PBS containing 0.1% BSA.

Cell staining and total cell numbers. The collected cells were stained for macrophages, neutrophils, T cells, and eosinophils using the following markers: CD11c (eBioscience, N418), Siglec F (BD Biosciences, E50-2440), Ly6G (eBioscience, 1A8), CD11b (eBioscience, M1/70), F4/80 (eBioscience, BM8), CD3 (eBioscience, 145-2C11), CD4 (eBioscience, RM4-5), CD44 (eBioscience, IM7), CD69 (eBioscience, H1.2F3). Absolute cell numbers were determined using AccuCount Particles (Spherotech). Flow cytometry data was collected on FACS Canto I (Becton Dickinson) and analysed with FlowJo (Treestar, Inc.).

Microscopy and histology. For haematoxylin and eosin and periodic-acid-Schiff (PAS) staining of lung sections, mice were perfused with PBS and a cannula inserted into the trachea. The lungs were gently inflated with 10% formalin at a constant fluid pressure at 25 cm. The trachea was tied off and the entire heart and lung were removed and placed in 10% formalin. Lungs were paraffin embedded, sectioned and stained by HistoTox Labs (Boulder). Additional lung sections were embedded and sectioned by Research Histology Core at University of Virginia and the immunohistochemical staining for IGF-1R, and cleaved caspase 3 was performed by the University of Virginia Biorepository and Tissue Research Facility. Approximately 6–10 images were taken per mice, with a total of 3–4 mice per group, and blindly scored by two independent scorers for inflammation, PAS staining, and cleaved caspase 3 positive cells. Immunofluorescence staining of lung sections was performed at the University of Virginia Cardiovascular Research Center Histological Services.

Airway hypersensitivity. Mice were anaesthetized and given a tracheotomy tube that delivered increasing concentrations of aerosolized methacholine. The

tracheotomy tube in turn was connected to the inspiratory and expiratory ports of a volume-cycled ventilator (flexiVent; SCIREQ Scientific). Airway resistance was measured at baseline and after each dose of methacholine³⁵.

Macrophage isolation. To obtain bone-marrow-derived macrophages, femurs were removed from 8-week-old mice and flushed with 5 ml of sterile PBS containing 5% FBS. The cell suspension was centrifuged, treated with red blood cell lysis buffer, washed, and then plated onto sterile Petri dishes in DMEM containing 10% L929 media, 10% FBS and 1% penicillin/streptomycin/glutamine. Media was replenished every 2 to 3 days and differentiated cells were used at day 6 after collection. Resident peritoneal macrophages were obtained by flushing the peritoneal cavity of mice with 10 ml of cold PBS containing 5% FBS. Collected cells were spun down, resuspended in X-VIVO 10 (Lonza) and plated at a concentration of 3×10^5 cells per well in a 24-well plate for IGF-1 secretion assays, and 5×10^5 per well in a 24-well plate for engulfment assays. Floating cells were washed the next day and remaining peritoneal macrophages were used 2 days after isolation. Alveolar macrophages were isolated by flushing the lungs with 1 ml of cold PBS instilled intratracheally (five flushes). Collected cells were centrifuged, resuspended in F12K media containing 10% FBS and 1% penicillin/streptomycin/glutamine, and seeded at 1×10^5 cells per well in a 48-well plate. All floating cells were washed away the next day and remaining cells were used in assays 2 days after isolation.

Microvesicle isolation. MH-S, mouse alveolar macrophages, or primary mouse alveolar macrophages were seeded. After adherence, the media was replaced with 0.22-µm-filtered media to remove any contaminating microvesicles. After an overnight incubation, the supernatant was collected and spun at 5,000g to remove cell debris and apoptotic bodies. The pellet was discarded and the resulting supernatant was filtered through a 0.8 µm filter and spun again at 17,000g. The pelleted microvesicles were then washed with HBSS and then spun again at 17,000g. For engulfment assays, microvesicles were stained with TAMRA for 20 min and then added to BEAS-2B cells for 90 min. For flow cytometry, purified particles were stained with CD11c, Siglec F, and Annexin V (BD Biosciences, 550475) and processed on ImageStreamX imaging flow cytometer (Amnis). Microvesicle size distribution was characterized using qNano (IZON Science) with a NP400 membrane and at least 500 particles were counted. Microvesicles were prepared for cryo-electron microscopy using standard methods and imaged on an FEI TF20 (ref. 36).

***In vivo* cytokine or apoptotic cell administration.** Mice were administered intranasally with 1 µg of recombinant mouse IL-4, IL-5, or IL-13 (eBiosciences), 1×10^6 apoptotic Jurkat cells, or PBS as control, for two consecutive days. On the third day, BAL fluid was recovered and centrifuged; the supernatant was stored at -80 °C for subsequent cytokine analysis.

Cytokine and IGF-1 analysis. IL-4, IL-5, IL-6, and CCL-11/Eotaxin-1 in the BAL fluid of mice that were sensitized with HDM (first regimen) were quantified by a multiplex Luminex performed by the University of Virginia Flow Cytometry Core Facility. Secretion of IGF-1 from J774 cells, peritoneal macrophages, and BAL fluids, as well as TSLP from BAL fluids, were measured by ELISA (R&D Systems).

***In vitro* cell systems.** LR73 (hamster fibroblasts), SVEC-40 (mouse endothelial cells), BEAS-2B (human bronchial epithelial cells, ATCC #CRL-9609), 16HBE14o-cells (human bronchial epithelial cells), MH-S (mouse alveolar macrophage cells ATCC #CRL-2019), Jurkat (human T cells) were either available in the laboratory or obtained from ATCC, with the latter tested for mycoplasma.

***In vitro* engulfment assay.** LR73, SVEC-40, BEAS-2B, and 16HBE14o-cells were seeded in a 24-well plate. Thymocytes were isolated from 4–6-week-old mice and induced to undergo apoptosis with dexamethasone. Thymocytes or prepared microvesicles were then stained with either CypHer5E (GE Healthcare, PA15401) or TAMRA (Invitrogen, C-1171). LR73 and SVEC-40 cells were incubated with apoptotic thymocytes at a 1:10 phagocyte to target ratio, BEAS-2B at a 1:5 phagocyte to target ratio, and 16HBE14o-cells at a 1:20 phagocyte to target ratio for 2 h. Mouse IGF-1 (Sigma-Aldrich), human IGF-1 (Sigma-Aldrich), human IGF-II (Sigma-Aldrich) were added to the phagocytes at the same time as addition of apoptotic targets. For IGFBP3 studies, IGFBP3 (Sigma-Aldrich) was added to media or supernatant from J774 cells for one h to allow IGFBP3 to bind to any available IGF-1, then the mix is added to phagocytes along with apoptotic targets. For all pharmacological studies, phagocytes were pre-incubated with the compounds listed below for one h before addition of apoptotic targets: cytochalasin D³⁷ (Sigma-Aldrich, C8273, 1 µM), latrunculin A³⁸ (Tocris, 3973, 150 nM), CK-666 (ref. 39) (Tocris, 3950, 25–100 µM), OSI-906 (ref. 40) (Selleckchem, S1091, 5–40 nM), NVP-AEW541 (ref. 41) (Selleckchem, S1034, 25–100 nM), Rapamycin (Sigma-Aldrich, R0395, 10 µM–1 mM), MK-2206 2HCl (Selleckchem, S1078, 10 nM–1 µM), U0126-EtOH (Selleckchem, S1102, 8 pM–5 nM), Wortmannin (Sigma-Aldrich, W1628, 50–200 nM), Y27632 (Calbiochem, 688000, 3.75–15 µM), or GSK269962 (Tocris, 4009, 80 nM–2 µM). Phagocytes were examined to ensure no gross morphological changes occurred owing to drug treatment. Targets were

then washed off three times with PBS, and the cells were dissociated from the plate with trypsin and the engulfment assessed by flow cytometry.

In vivo engulfment assay. CCSP-Cre/YFP⁺ mice were administered PBS or 1 µg of IGF-1 intranasally. One h later, 100 million CypHer5E-labelled apoptotic thymocytes with or without IGF-1 were injected intranasally for 3.5 h. The BAL fluid was collected and the lungs excised, minced, and digested into a single-cell suspension. The cells were then stained with appropriate markers: CD11c and Siglec F for alveolar macrophage markers and EpCam (along with YFP expression) for airway epithelial cells and analysed by flow cytometry to assess apoptotic cell uptake by the airway epithelial cells and the alveolar macrophages.

Liposome construction. Liposomes were prepared by dissolving the lipids (phosphatidylserine, dioleoyl phosphatidylcholine, cholesterol and the lipid DiD dye) in chloroform, evaporating chloroform under flow of argon gas in a glass vial, then subjecting the lipid layer to overnight lyophilization to remove traces of organic solvent. Then normal saline was added for hydration, and intense vortexing was performed to prepare multilamellar vesicles (MLV). Liposomes were repeatedly filtered through a 0.2-µm Nuclepore polycarbonate filter to prepare smaller particles. Particle size was verified by dynamic light scattering using Nicomp 370.

Immunoblotting. LR73, J774 or BEAS-2B cells were seeded in a 60-mm dish at a concentration of 5×10^5 . Cells were serum-starved for 6 h and then stimulated with 100 ng ml⁻¹ of IGF-1 for various time points. Cells were lysed in RIPA buffer and used in western blots. The blots were probed for phospho-Erk1/2 (Cell Signaling Technology, 4370), phospho-Akt (Cell Signaling Technology, 4060), phospho-IGF-1R (Cell Signaling, 3024), total Erk2 (Santa Cruz Biotechnology, sc-154-G), total Akt (Cell Signaling Technology, 4691), total IGF-1R (Cell Signaling, 9750), and anti-B-actin-HRP (Sigma-Aldrich, A3854) followed by chemiluminescence detection.

Quantitative RT-PCR. Total RNA was extracted from cells using Quick-RNA Miniprep Kit (Zymo Research) or RNeasy Mini Kit (Qiagen) and cDNA was synthesized using QuantiTect Reverse Transcription Kit (Qiagen) according to manufacturers' instructions. Quantitative gene expression for mouse *Igf1*, human *TSLP*, *CSF2*, *IL6*, *IL8* or housekeeping human or mouse *Hprt* was performed using Taqman probes (Applied Biosystems) using StepOnePlus Real Time PCR System (ABI).

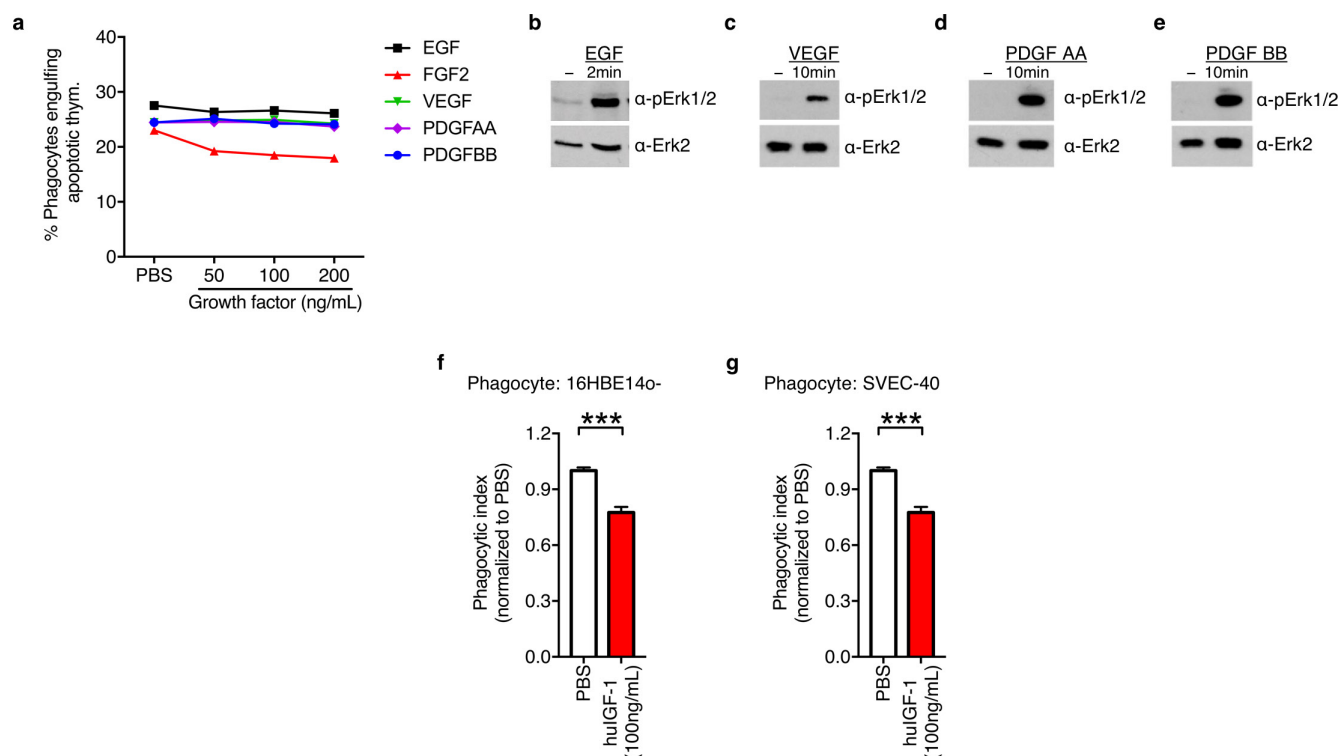
RNA-seq. BEAS-2B cells were treated with HDM and alveolar macrophage-derived microvesicles for 3 h. Total RNA was extracted and an mRNA library was prepared using Illumina TruSeq platform and followed by transcriptome sequencing using an Illumina NextSeq 500 cartridge. Four independent experiments were sequenced. Rv3.2.2 was used for graphical and statistical analysis and the R package DESeq2 was used for differential gene expression analysis of RNA-seq data. RNA-seq analysis was performed by the UVA Bioinformatics Core.

Code availability. R code used for bioinformatics analysis and heat map generation is available upon request.

Data availability. RNA-seq data have been deposited in the Gene Expression Omnibus under accession code GSE89049.

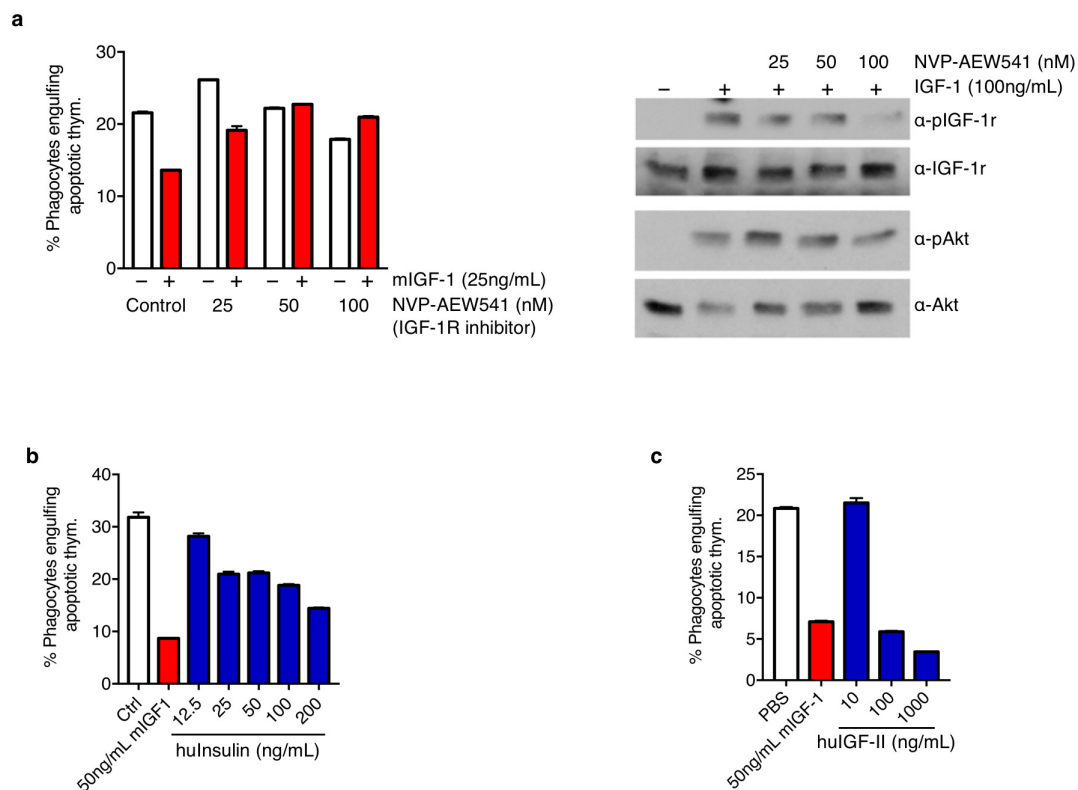
Statistical analysis. Statistical significance was determined using GraphPad Prism 5 or 6 using unpaired Student's two-tailed *t*-test, one-sample *t*-test, one-way ANOVA or two-way ANOVA, according to test requirements. Variance was similar between groups. No inclusion/exclusion criteria were pre-established. Grubbs' Outlier Test was used to determine outliers, which were excluded from final analysis. A *P* value of <0.05 (indicated by one asterisk), <0.01 (indicated by two asterisks), or <0.001 (indicated by three asterisks) were considered significant.

31. Perl, A. K., Zhang, L. & Whitsett, J. A. Conditional expression of genes in the respiratory epithelium in transgenic mice: cautionary notes and toward building a better mouse trap. *Am. J. Respir. Cell Mol. Biol.* **40**, 1–3 (2009).
32. Dietrich, P., Dragatsis, I., Xuan, S., Zeitlin, S. & Efstratiadis, A. Conditional mutagenesis in mice with heat shock promoter-driven cre transgenes. *Mamm. Genome* **11**, 196–205 (2000).
33. Liu, J. L., Yakar, S. & LeRoith, D. Conditional knockout of mouse insulin-like growth factor-1 gene using the Cre/loxP system. *Proc. Soc. Exp. Biol. Med.* **223**, 344–351 (2000).
34. Liu, J. L. *et al.* Insulin-like growth factor-I affects perinatal lethality and postnatal development in a gene dosage-dependent manner: manipulation using the Cre/loxP system in transgenic mice. *Mol. Endocrinol.* **12**, 1452–1462 (1998).
35. Hahn, Y. S. *et al.* Vγ4⁺ γδ T cells regulate airway hyperreactivity to methacholine in ovalbumin-sensitized and challenged mice. *J. Immunol.* **171**, 3170–3178 (2003).
36. Yeager, M., Berriman, J. A., Baker, T. S. & Bellamy, A. R. Three-dimensional structure of the rotavirus haemagglutinin VP4 by cryo-electron microscopy and difference map analysis. *EMBO J.* **13**, 1011–1018 (1994).
37. Casella, J. F., Flanagan, M. D. & Lin, S. Cytochalasin D inhibits actin polymerization and induces depolymerization of actin filaments formed during platelet shape change. *Nature* **293**, 302–305 (1981).
38. Yarmola, E. G., Somasundaram, T., Boring, T. A., Spector, I. & Bubb, M. R. Actin-latrunculin A structure and function. Differential modulation of actin-binding protein function by latrunculin A. *J. Biol. Chem.* **275**, 28120–28127 (2000).
39. Hetrick, B., Han, M. S., Helgeson, L. A. & Nolen, B. J. Small molecules CK-666 and CK-869 inhibit actin-related protein 2/3 complex by blocking an activating conformational change. *Chem. Biol.* **20**, 701–712 (2013).
40. Mulvihill, M. J. *et al.* Discovery of OSI-906: a selective and orally efficacious dual inhibitor of the IGF-1 receptor and insulin receptor. *Future Med. Chem.* **1**, 1153–1171 (2009).
41. García-Echeverría, C. *et al.* In vivo antitumor activity of NVP-AEW541—a novel, potent, and selective inhibitor of the IGF-IR kinase. *Cancer Cell* **5**, 231–239 (2004).



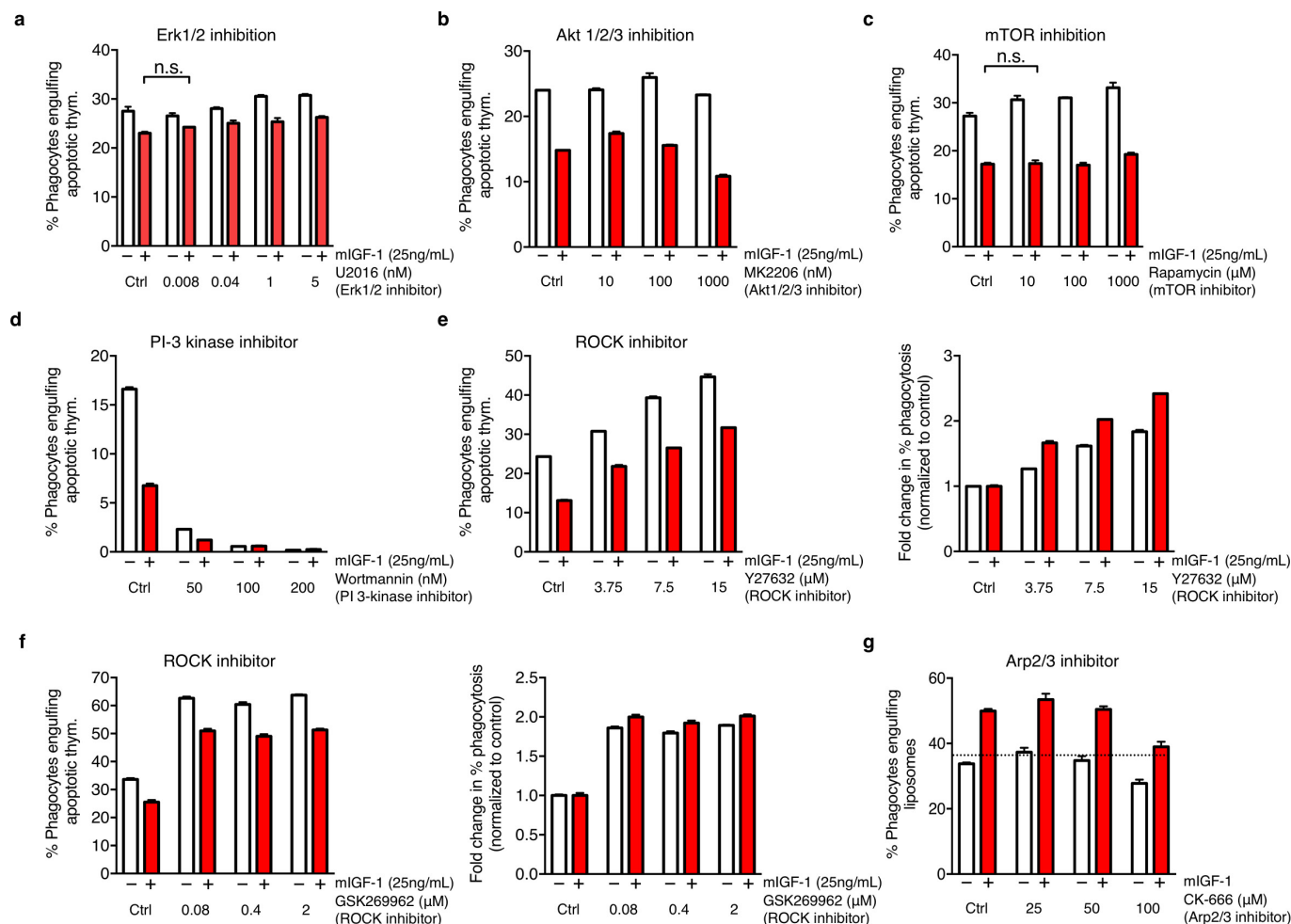
Extended Data Figure 1 | IGF-1, but not EGF, VEGF, PDGF-AA/-BB, suppresses phagocytosis of apoptotic cells in non-professional phagocytes. **a**, Representative engulfment assay in which LR73 cells were treated with indicated growth factors at increasing concentrations and assessed for engulfment of apoptotic thymocytes ($n = 3$). **b–e**, Serum-starved LR73 cells were stimulated with 100 ng ml^{-1} of

specified growth factors for indicated time and the phosphorylation of Erk1/2 was determined by immunoblotting ($n = 2$). **f**, **g**, Representative engulfment assay in which the uptake of apoptotic thymocytes by 16HBE14o- human airway epithelial cells (**f**) or SVEC-40 endothelial cells (**g**) is dampened by IGF-1 treatment ($n = 3$). Error bars represent s.d., P value of <0.001 (***).



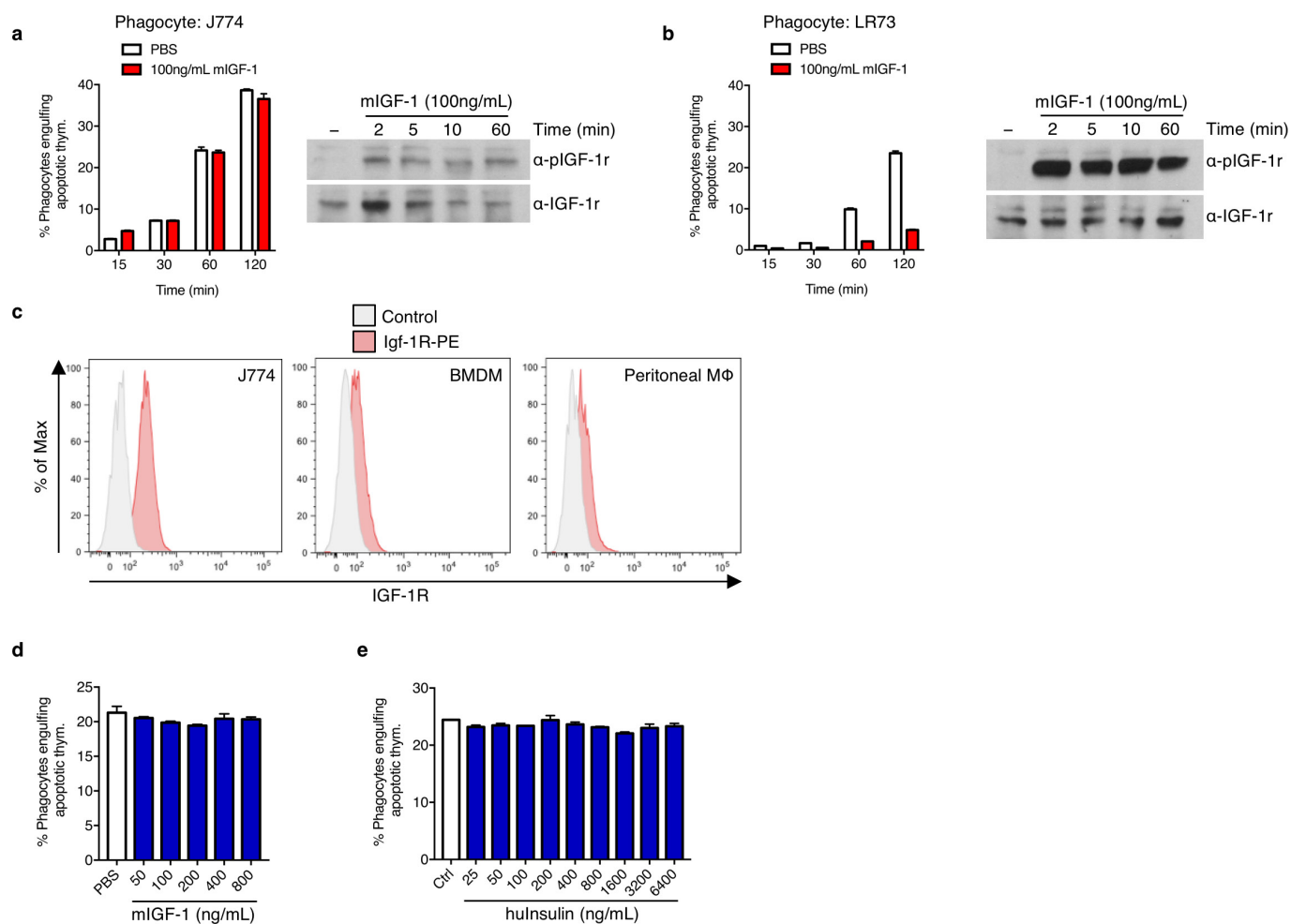
Extended Data Figure 2 | Insulin and IGF-II also decrease apoptotic cell engulfment, similar to IGF-1 that is reversed by treatment with NVP-AEW541. **a**, Left, engulfment of apoptotic thymocytes by LR73 cells treated with various doses of NVP-AEW541, a small molecule inhibitor of IGF-1R ($n = 3$). Right, representative immunoblot of LR73 cells stimulated

with IGF-1 and treated with increasing doses of NVP-AEW541 ($n = 2$). **b**, **c**, Engulfment of apoptotic thymocytes by LR73 cells treated with the indicated concentrations of human insulin (**b**) and human IGF-II (**c**) ($n = 2$). Error bars represent s.d.



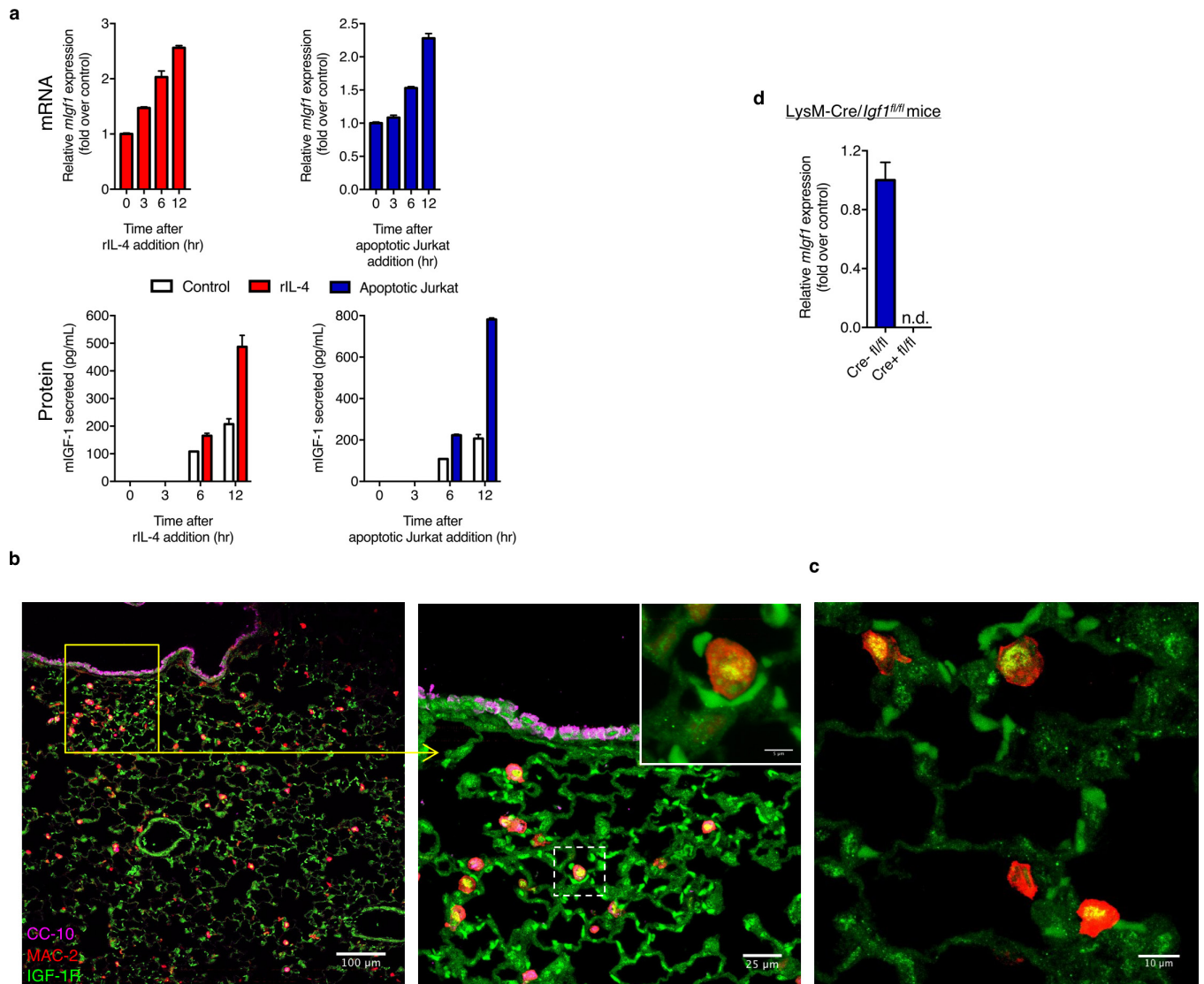
Extended Data Figure 3 | Blocking canonical signalling intermediates downstream of IGF-1 receptor signalling, Rho-kinase- or Arp2/3-mediated functions, does not reverse the IGF-1 mediated engulfment modulation. a–f, Engulfment of apoptotic thymocytes by LR73 cells treated with U2016 (Erk1/2 inhibitor) (a), MK-2206 (Akt1/2/3 inhibitor) (b), Rapamycin (mTOR inhibitor) (c), or Wortmannin (PI 3-Kinase inhibitor) (d), Rho-kinase inhibitors Y27632 (e) or GSK269962 (f) in the presence or absence of IGF-1 ($n = 2-3$). Initially, it appeared that inhibition of Rho-kinase was able to partially rescue IGF-1 induced engulfment suppression. However, as Rho-kinase inhibition

basally increases phagocytosis of apoptotic cells (consistent with what has been previously reported), we normalized the change in phagocytosis for each inhibitor concentration to the appropriate control (right panel). After normalizing, we observed that Rho-kinase inhibition did not increase apoptotic cell uptake in LR73 cells in the presence of IGF-1 more than the increase observed basally owing to Rho-kinase inhibition. Thus, inhibition of Rho-kinase does not appear to rescue IGF-1-induced engulfment suppression. g, LR73 cells were treated with CK-666 and then assessed for uptake of liposomes in the presence of IGF-1. Data represented as mean \pm s.d.



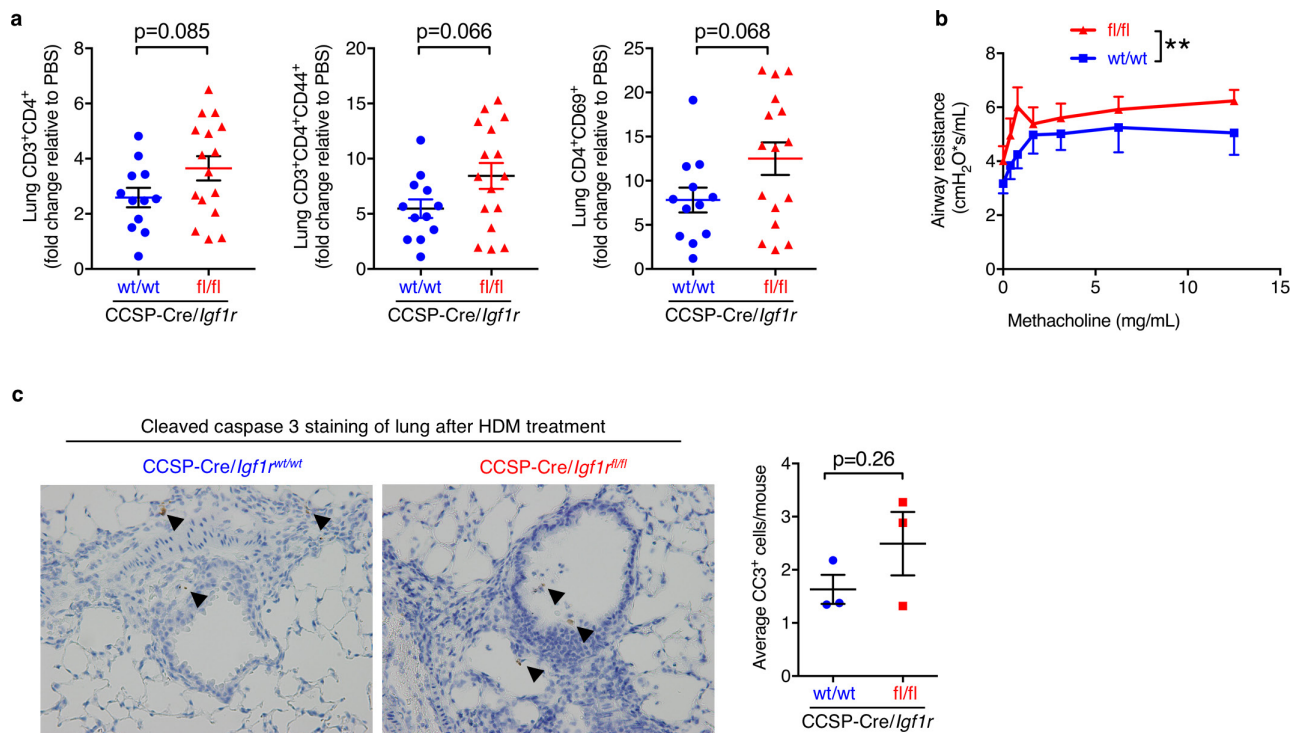
Extended Data Figure 4 | Macrophages express IGF-1R and phosphorylate IGF-1R upon IGF-1 stimulation, but engulf apoptotic cells at normal capacity when exposed to IGF-1 or insulin. **a, b,** J774 cells (**a**) or LR73 cells (**b**) treated with 100 ng ml⁻¹ mouse IGF-1 were assessed for their ability to engulf apoptotic thymocytes, or serum-starved for 6 h and stimulated with 100 ng ml⁻¹ mouse IGF-1 and assessed for

phosphorylation of IGF-1R by western blot. **c,** Flow cytometry histograms of IGF-1R expression on J774 cells (left), bone-marrow-derived macrophages (middle), and peritoneal macrophages (right) ($n = 3-4$). **d, e,** IC-21 cells treated with indicated concentrations of mouse IGF-1 (**d**) or human insulin (**e**) were assessed for their ability to engulf apoptotic thymocytes ($n = 2-3$). Error bars represent s.d.



Extended Data Figure 5 | Production of IGF-1 by peritoneal macrophages after apoptotic cell or IL-4 stimulation correlates with new transcription. **a**, Peritoneal macrophages were either untreated, stimulated with rIL-4 or apoptotic Jurkat cells and *Igf1* mRNA (top panels) and IGF-1 protein in the supernatant (bottom panels) were assessed in a time course (n.d., not detected) ($n = 3$). Data represented as mean \pm s.d.

b, c, Lung sections from wild-type mice were stained with antibodies against alveolar macrophages (Mac-2), airway epithelial cells (CC-10), and IGF-1R. **d**, Alveolar macrophages isolated from LysM-Cre/*Igf1*^{fl/fl} and littermate controls were assessed for *Igf1* mRNA expression ($n = 2$ per group, data represented as mean \pm s.e.m.).

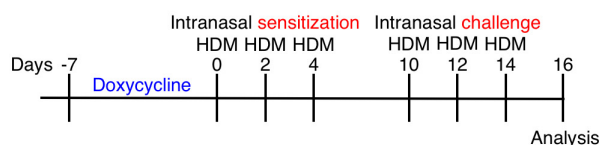


Extended Data Figure 6 | CCSP-Cre/Igf1r^{fl/fl} mice exposed to HDM have greater airway resistance and show a trend towards greater immune cell infiltration in the lungs and more apoptotic cells.
a, Total cell counts of lung CD3⁺CD4⁺ T cells (left), CD3⁺CD4⁺CD44⁺ T cells (middle), and CD3⁺CD4⁺CD69⁺ T cells (right panel) in the lungs of CCSP-Cre/Igf1r^{+/+} and Igf1r^{fl/fl} mice given the full HDM course.
b, Airway hyper-responsiveness to methacholine (another measure of

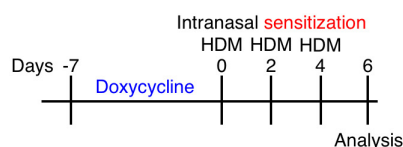
allergen sensitivity) in the CCSP-Cre/Igf1r^{fl/fl} mice compared to control CCSP-Cre/Igf1r^{+/+} mice treated with HDM ($n=6-8$ mice per group).
c, Representative histology images of cleaved caspase (CC3) staining in lung sections of mice given the full HDM course. Black arrowheads indicate positive staining. Average CC3-positive cells per mouse are quantified on the right ($n=3$ per group). Data represented as mean \pm s.e.m. P value of <0.01 (**).

a

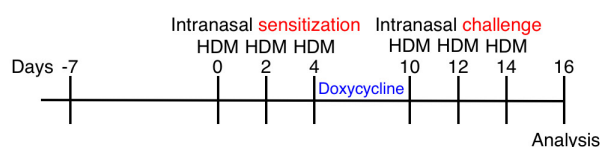
Full Course



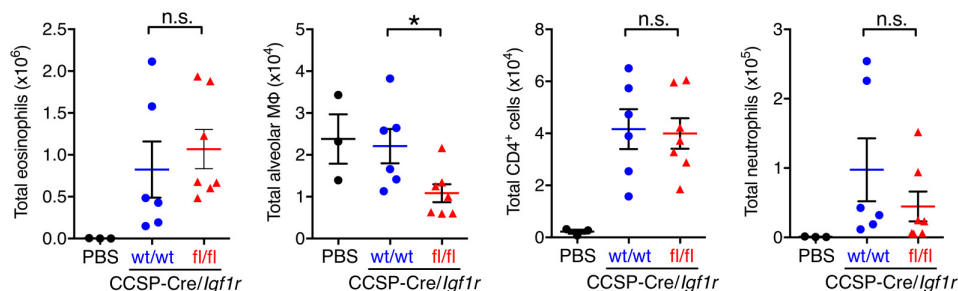
Regimen #1: Sensitization Phase



Regimen #2: Challenge Phase

**b**

Bronchoalveolar lavage (day 16)

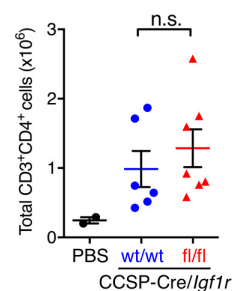


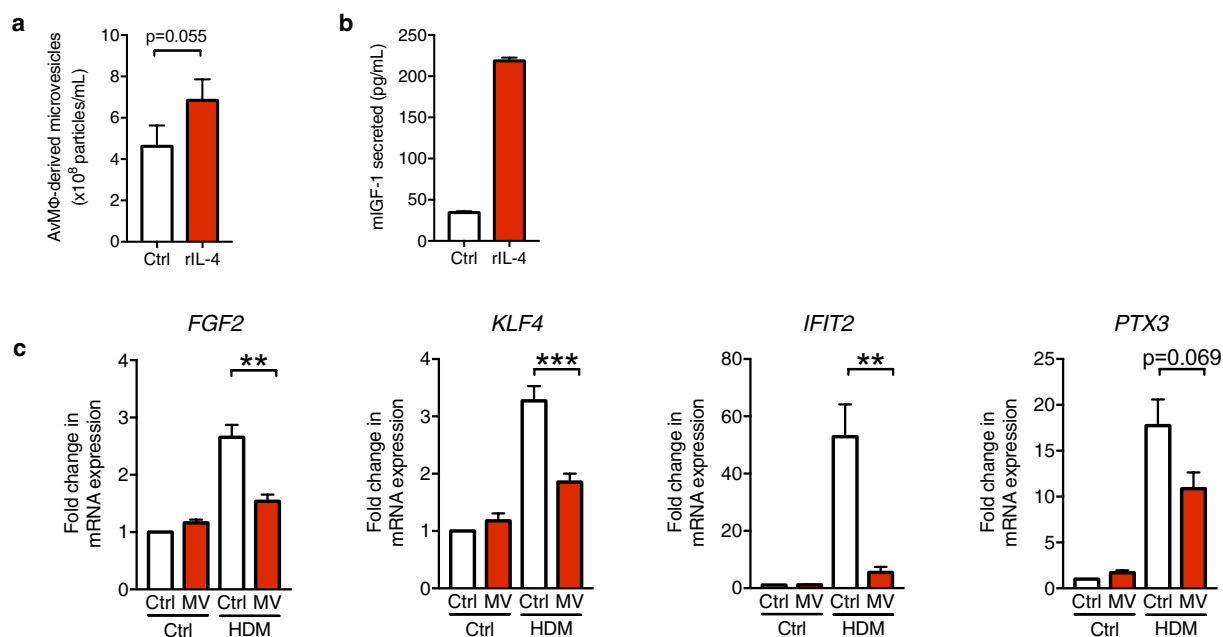
Extended Data Figure 7 | Schematic IGF-1R deletion during the sensitization versus challenge phases of HDM administration, and the response of CCSP-Cre/Igf1r^{+/+} and Igf1r^{fl/fl} mice in the second regime (the challenge phase). **a**, Schematic describing the different time courses for *Igf1r* deletion from Club cells (induced via administration of doxycycline) and for the allergen HDM exposure. **b**, Total cell counts of

various populations in the BAL fluid of CCSP-Cre/Igf1r^{+/+} and CCSP-Cre/Igf1r^{fl/fl} mice given HDM according to the second regimen. **c**, Total cell counts of CD3⁺CD4⁺ T cells of draining lymph nodes of CCSP-Cre/Igf1r^{+/+} and CCSP-Cre/Igf1r^{fl/fl} mice given HDM according to the second regimen. Data represented as mean \pm s.e.m. NS, not significant. *P* value of <0.05 (*).

c

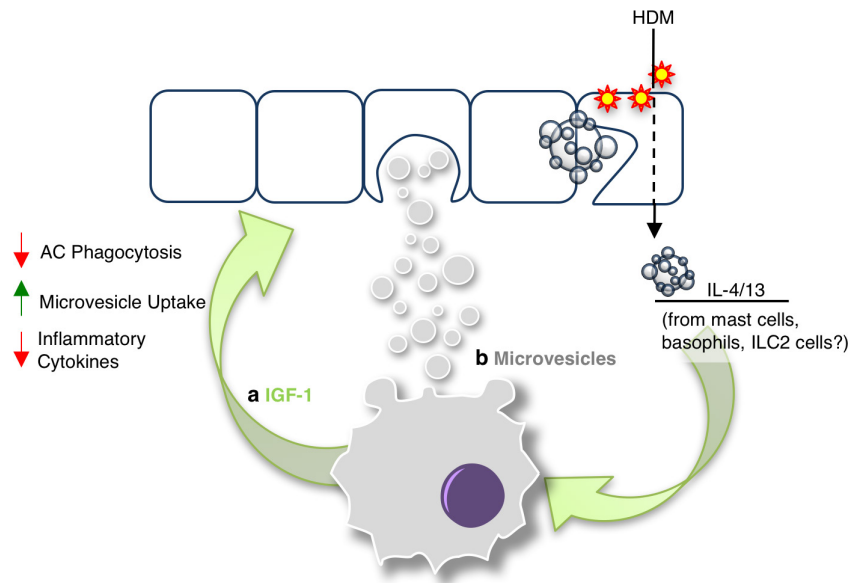
Draining lymph nodes (day 16)





Extended Data Figure 8 | Alveolar-macrophage-derived microvesicles suppress gene expression in lung epithelial cells exposed to HDM extract. **a**, Microvesicles (MV) were collected from either control or IL-4-treated MH-S alveolar macrophages and then counted using qNano ($n = 3$). **b**, Supernatants from IL-4-treated MH-S macrophages

were assessed for IGF-1 secretion. **c**, BEAS-2B cells were treated with HDM either in the presence or absence of alveolar-macrophage-derived microvesicles for 3 h and then assessed for expression of *FGF2*, *KLF4*, *IFIT2*, and *PTX3* ($n = 6$). Data represented as mean \pm s.e.m. P value of <0.01 (**), or <0.001 (***)



Extended Data Figure 9 | Model for alveolar macrophage regulation of airway epithelial cells, with respect to particle uptake and the response to allergens, through IGF-1 and microvesicles. Exposure of airways to allergens, such as HDM, can cause apoptotic cell death as well as IL-4 and IL-13 production, from mast cells and type-2 innate lymphoid cells (ILC2s). These cytokines, along with apoptotic cells, trigger alveolar

macrophages to produce IGF-1. The released IGF-1 (a) then acts on the airway epithelium to elicit two actions: first, to decrease the uptake of apoptotic cells, and second, to enhance the uptake of macrophage-derived microvesicles. These microvesicles (b) dampen inflammatory cytokine production through airway epithelial cells.

Control of mitochondrial function and cell growth by the atypical cadherin Fat1

Longyue L. Cao^{1,2*}, Dario F. Riascos-Bernal^{1,2*}, Prameladevi Chinnasamy^{1,2}, Charlene M. Dunaway^{1,2}, Rong Hou^{1,2†}, Mario A. Pujato^{3†}, Brian P. O'Rourke⁴, Veronika Miskolci⁵, Liang Guo⁶, Louis Hodgson^{5,7}, Andras Fiser³ & Nicholas E. S. Sibinga^{1,2}

Mitochondrial products such as ATP, reactive oxygen species, and aspartate are key regulators of cellular metabolism and growth. Abnormal mitochondrial function compromises integrated growth-related processes such as development and tissue repair^{1,2}, as well as homeostatic mechanisms that counteract ageing and neurodegeneration³, cardiovascular disease^{4,5}, and cancer^{6,7}. Physiologic mechanisms that control mitochondrial activity in such settings remain incompletely understood. Here we show that the atypical Fat1 cadherin acts as a molecular 'brake' on mitochondrial respiration that regulates vascular smooth muscle cell (SMC) proliferation after arterial injury. Fragments of Fat1 accumulate in SMC mitochondria, and the Fat1 intracellular domain interacts with multiple mitochondrial proteins, including critical factors associated with the inner mitochondrial membrane. SMCs lacking Fat1 (*Fat1*^{KO}) grow faster, consume more oxygen for ATP production, and contain more aspartate. Notably, expression in *Fat1*^{KO} cells of a modified Fat1 intracellular domain that localizes exclusively to mitochondria largely normalizes oxygen consumption, and the growth advantage of these cells can be suppressed by inhibition of mitochondrial respiration, which suggest that a Fat1-mediated growth control mechanism is intrinsic to mitochondria. Consistent with this idea, Fat1 species associate with multiple respiratory complexes, and Fat1 deletion both increases the activity of complexes

I and II and promotes the formation of complex-I-containing supercomplexes. *In vivo*, Fat1 is expressed in injured human and mouse arteries, and inactivation of SMC Fat1 in mice potentiates the response to vascular damage, with markedly increased medial hyperplasia and neointimal growth, and evidence of higher SMC mitochondrial respiration. These studies suggest that Fat1 controls mitochondrial activity to restrain cell growth during the reparative, proliferative state induced by vascular injury. Given recent reports linking Fat1 to cancer, abnormal kidney and muscle development, and neuropsychiatric disease^{8–13}, this Fat1 function may have importance in other settings of altered cell growth and metabolism.

Proliferation of SMCs is essential for repair of injured arteries, but in excess can cause vascular obstruction leading to tissue ischaemia and/or infarction¹⁴. SMC expression of the atypical cadherin Fat1 is induced by both growth factor exposure and vascular injury; however, knockdown of Fat1 enhances SMC proliferation, and the Fat1 intracellular domain (Fat1 ICD), expressed as a chimaeric protein fused to the interleukin 2 receptor α -chain (denoted hereafter as Fat1-IL-2R), is sufficient to inhibit growth¹⁵. In addition, recent reports link *FAT1* silencing and/or mutation to a spectrum of cancers^{8,9}. To understand how Fat1 controls proliferation, we used tandem affinity purification (TAP) followed by peptide mass fingerprinting using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (TAP-MS) to identify

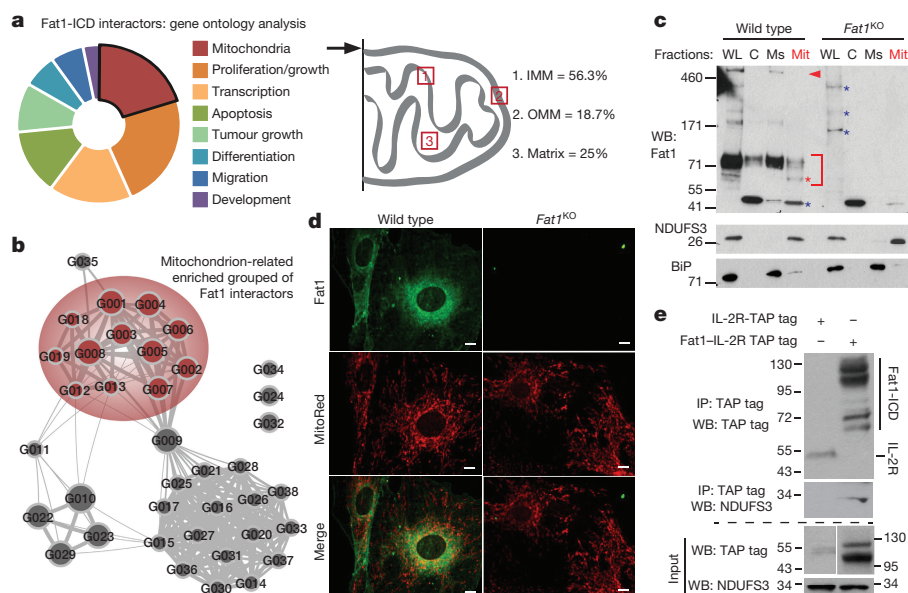


Figure 1 | Fat1 fragments localize to SMC mitochondria and interact with inner mitochondrial membrane proteins. a, Analysis of 30 top-ranked TAP-MS-validated Fat1 ICD interactors. IMM and OMM, inner and outer mitochondrial membrane, respectively (see Extended Data Table 1). **b**, Mitochondrial cluster in bioinformatic analysis of TAP-MS data (see Extended Data Fig. 1a). **c**, Fractionation of mouse aortic SMCs, followed by SDS-PAGE and immunoblotting. C, cytoplasmic; Mit, mitochondrial; Ms, microsomal; WL, whole-cell lysate. Arrowhead, full-length Fat1; bracket, Fat1 ICD species; red asterisk, mitochondrial-specific Fat1 ICD fragments; blue asterisk, non-specific signal. **d**, Mouse SMC confocal imaging. Scale bar, 10 μ m. **e**, Co-immunoprecipitation of Fat1 ICD and NDUFS3 in 293T cells. For gel source data, see Supplementary Fig. 1.

¹Wilf Family Cardiovascular Research Institute, Department of Medicine (Cardiology), Albert Einstein College of Medicine, Bronx, New York 10461, USA. ²Department of Developmental and Molecular Biology, Albert Einstein College of Medicine, Bronx, New York 10461, USA. ³Department of Systems & Computational Biology, Albert Einstein College of Medicine, Bronx, New York 10461, USA. ⁴Department of Physiology and Biophysics, Albert Einstein College of Medicine, Bronx, New York 10461, USA. ⁵Department of Anatomy and Structural Biology, Albert Einstein College of Medicine, Bronx, New York 10461, USA. ⁶CVPath Institute, Gaithersburg, Maryland 20878, USA. ⁷Gruss-Lipper Biophotonics Center, Albert Einstein College of Medicine, Bronx, New York 10461, USA. [†]Present addresses: Department of Internal Medicine, Division of Infectious Diseases, Allergy, and Immunology, St. Louis University Hospital, St. Louis, Missouri 63104, USA (R.H.); Center for Autoimmune Genomics and Etiology, Cincinnati Children's Hospital Medical Center, 3333 Burnet Avenue, Cincinnati, Ohio 45229, USA (M.A.P.).

*These authors contributed equally to this work.

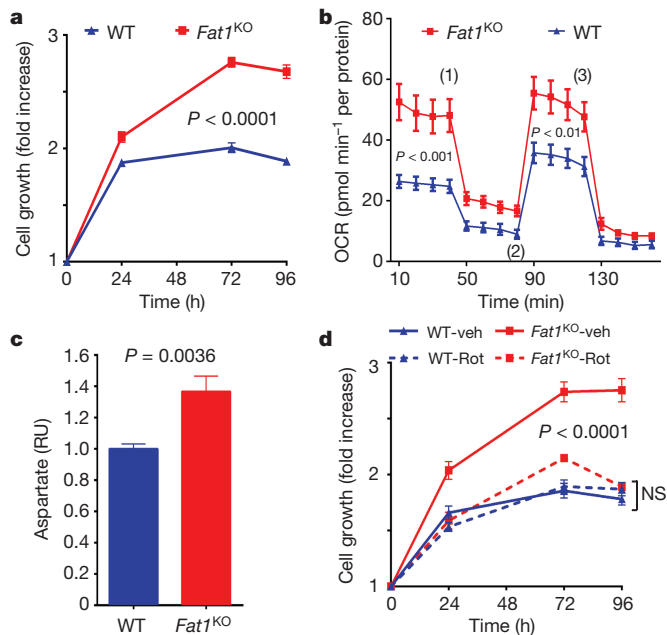


Figure 2 | Fat1 suppresses SMC growth by restraining mitochondrial respiration. **a**, Growth of mouse aortic SMCs. $n = 6$. **b**, OCR of SMCs at baseline and in response to oligomycin (1), carbonyl cyanide-*p*-trifluoromethoxyphenylhydrazone (FCCP, an uncoupling agent that disrupts ATP synthesis) (2), and rotenone (3). $n = 10$. **c**, SMC aspartate content. RU, relative units. $n = 7$, analysed by two-tailed *t*-test. **d**, SMC growth in response to 1 μ M rotenone (rot) or vehicle (veh). $n = 3$. Data assessed by two-way ANOVA (**a**, **b**, **d**). All data shown as mean \pm s.e.m.

proteins that interact with the Fat1 ICD. We noted that a surprisingly large fraction of candidate interactors associated with mitochondria, including several proteins that localize to the inner mitochondrial membrane (Fig. 1a, Extended Data Table 1); these observations were replicated in an unbiased independent bioinformatic analysis of the TAP-MS dataset (Fig. 1b, Extended Data Fig. 1a).

We used a conditionally targeted mouse *Fat1* allele (*Fat1^{loxP}*)¹¹ and Cre-mediated recombination to inactivate *Fat1* expression in SMCs (Extended Data Fig. 1b). Cellular fractionation showed *Fat1* fragments in the mitochondrial sample: immunoblotting using a *Fat1* C-terminus-directed antiserum¹⁵ identified multiple specific signals between ~60–100 kDa, including an ~60 kDa band seen only in whole-cell and mitochondrial fractions (Fig. 1c). Confocal imaging of these cells showed *Fat1* and mitochondrial co-localization (Fig. 1d). Identification of multiple inner mitochondrial membrane proteins in our screen (Fig. 1a, Extended Data Fig. 1a, Extended Data Table 1) suggests that *Fat1* may associate with respiratory complexes; indeed, traditional co-immunoprecipitation studies confirmed interaction of the *Fat1* ICD with NDUF3, a complex-I-specific subunit (Fig. 1e), and with prohibitin, an important regulator of complex I stability and function¹⁶ (Extended Data Fig. 1c).

We then assessed growth and mitochondrial function. Resazurin (AlamarBlue) growth/viability assays were validated across a range of input cell numbers, and performed similarly for both wild-type and *Fat1^{KO}* cells (Extended Data Fig. 2a). Evidence of higher proliferation in SMCs lacking *Fat1* (Fig. 2a, Extended Data Fig. 2b) was normalized by restoring *Fat1* expression with *Fat1*-IL-2R (Extended Data Fig. 2b). In assays of mitochondrial respiration using extracellular flux (Seahorse) analysis, *Fat1^{KO}* SMCs exhibited higher baseline and maximal oxygen consumption rates (OCRs) compared to wild type (Fig. 2b, Extended Data Fig. 2c), with equal coupling efficiency (Extended Data Fig. 2d). Coupling efficiency reflects how effectively respiratory activity is translated into ATP production¹⁷, so higher ATP-linked oxygen consumption in *Fat1^{KO}* SMCs indicates that these cells

generate more ATP. Net ATP levels in wild-type and *Fat1^{KO}* SMCs were similar (Extended Data Fig. 2e), which in turn suggests that consumption of ATP in *Fat1^{KO}* SMCs is also increased. Reactive oxygen species, which also result from mitochondrial respiration, were present at similar levels under standard conditions, but were higher in *Fat1^{KO}* cells after exposure to H_2O_2 (Extended Data Fig. 2f), suggesting a partially depleted capacity to buffer or inactivate reactive oxygen species when *Fat1* is absent. Aspartate, a metabolite essential for macromolecular synthesis and cell proliferation, and a third key product of respiration^{18,19}, was present at higher levels in *Fat1^{KO}* SMCs (Fig. 2c). Mitochondrial respiration is essential for cell growth^{18,19}, including that of SMCs²⁰, so we tested its role in the increased growth of *Fat1^{KO}* cells. Limiting respiration by either pharmacologic or genetic interference targeting complex I did not compromise wild-type cell growth, but suppressed the growth advantage of *Fat1^{KO}* cells (Fig. 2d, Extended Data Fig. 3a–c). These studies suggest that physiologic *Fat1* expression modulates SMC growth by restraining mitochondrial respiration, and that further limits on respiration in wild-type cells have no repercussion. By contrast, when *Fat1* is absent, the corresponding loss of this restraint increases SMC growth. Overall, these *in vitro* studies identify *Fat1* as an important suppressor of SMC growth, and show that potentiation of growth in the absence of *Fat1* depends on increased respiration.

Respiration can be affected by changes in mitochondrial structure, mass, and/or dynamics, but mitochondria from wild-type and *Fat1^{KO}* SMCs did not exhibit major ultrastructural differences by electron microscopy (Extended Data Fig. 4a–f). Likewise, we found no evidence for changes in (1) mitochondrial mass, on the basis of the levels of several respiratory chain subunits (Extended Data Fig. 4g), (2) mitochondrial biogenesis and autophagy, on the basis of expression of regulators of these respective processes²¹ (Extended Data Fig. 4h–j), or (3) mitochondrial dynamics, as reflected in fusion or fission gene expression (Extended Data Fig. 4h, k). Failing to identify differences

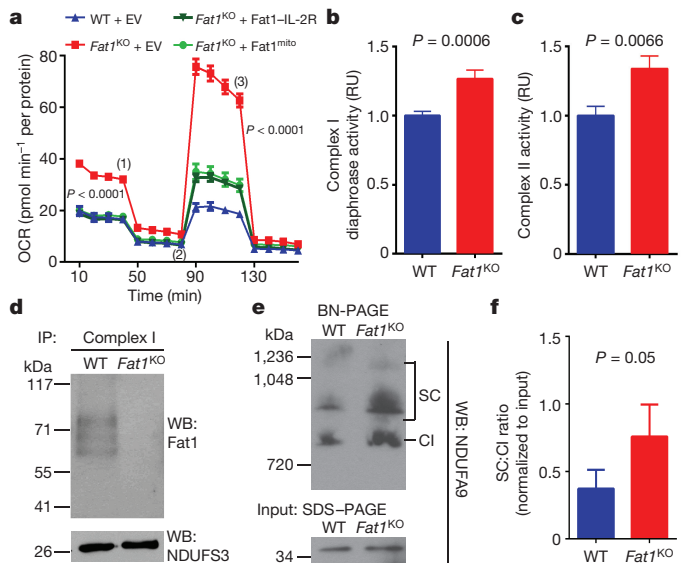


Figure 3 | The *Fat1* ICD limits complex I and II activities, and complex-I-containing supercomplex formation in SMCs. **a**, OCR after introducing *Fat1*-IL-2R or mitochondria-targeted *Fat1* ICD (*Fat1^{mito}*) into *Fat1^{KO}* SMCs. (1), oligomycin; (2), FCCP; (3), rotenone; EV, empty vector. $n = 15$, analysed by one-way ANOVA. **b**, NADH oxidation by immunocaptured complex I from SMC lysates, expressed as fold-increase from wild type. Wild type $n = 13$, *Fat1^{KO}* $n = 11$. **c**, Ubiquinol production by immunocaptured complex II. $n = 16$. **d**, SDS-PAGE and immunoblotting of proteins eluted from immunocaptured complex I. **e**, Native complex I (CI) and supercomplexes (SC) containing complex I in mouse SMC mitochondria by blue native (BN)-PAGE analysis. **f**, Quantification of SC to CI ratio from BN-PAGE analyses. $n = 5$. RU, relative units. Data analysed by two-tailed *t*-test (**b**, **c**, **f**). All data shown as mean \pm s.e.m. For gel source data, see Supplementary Fig. 1.

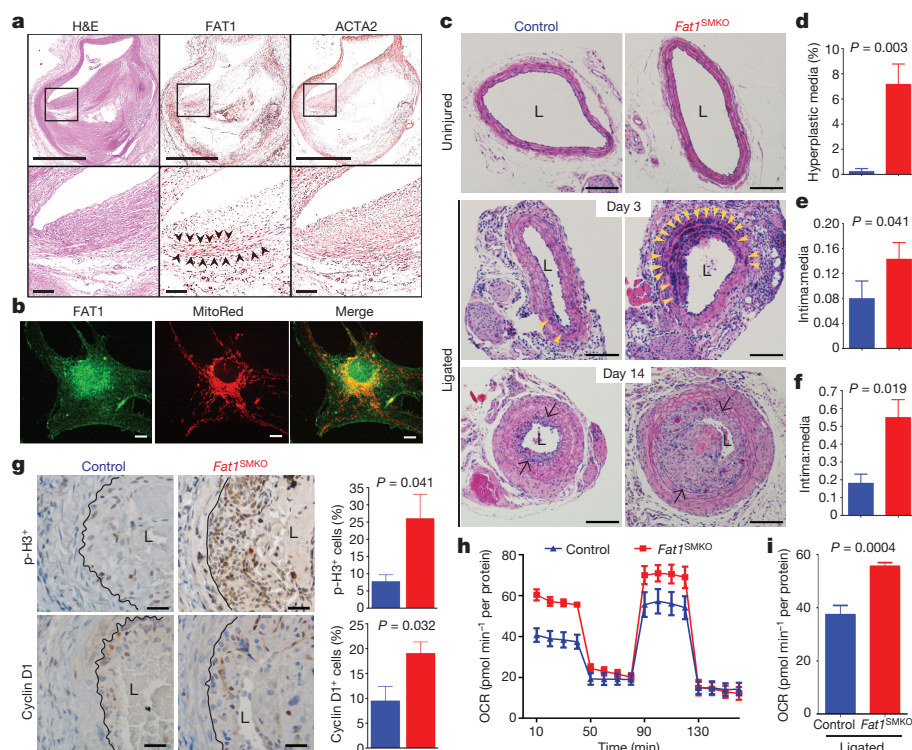


Figure 4 | Fat1 restrains SMC growth and mitochondrial respiration after vascular injury. **a**, Haematoxylin and eosin (H&E) stain and FAT1 and ACTA2 immunohistochemistry in restenotic human coronary arteries. Black squares are magnified in lower panels. Arrowheads, regions of co-expression. Scale bars, 1 mm (upper panels) and 100 μ m (lower panels). **b**, Human SMC confocal imaging. Scale bar, 10 μ m. **c**, H&E-stained mouse carotid arteries. Arrowheads, medial hyperplasia. Arrows, internal elastic lamina. L, lumen. Scale bar, 100 μ m. **d**, Hyperplasia as

percentage of total medial area. **e, f**, Intima:media ratio 3 and 14 days post-injury, respectively. $n = 5$ for control, $n = 8$ for *Fat1*^{SMKO} (**d–f**). **g**, Phospho-histone H3 (p-H3) and cyclin D1 14 days after injury. Graphs, percentage of positive (brown) versus total neointimal cells. $n = 6$ for p-H3, $n = 5$ for cyclin D1. Black lines mark the internal elastic lamina. Scale bar, 25 μ m. **h**, OCR of SMCs from ligated arteries, 3 days after injury. **i**, Basal respiration from **h**. $n = 7$, analysed by two-tailed *t*-test. Data analysed by two-tailed Mann–Whitney test (**d–g**). All data shown as mean \pm s.e.m.

in mitochondrial structure, mass, or turnover that could explain the increased respiration in *Fat1*^{KO} SMCs, we surmised that loss of Fat1 might directly affect respiratory chain activity.

To assess this possibility, we tested if the Fat1 ICD was sufficient to control respiration. Expression of Fat1–IL-2R—which inhibits SMC growth¹⁵—limited respiration in *Fat1*^{KO} SMCs (Fig. 3a), and yielded fragments in the mitochondrial fraction similar in size to endogenous mitochondrial Fat1 species (Extended Data Fig. 5a), but this fusion protein may also localize to other subcellular compartments. We generated and validated a mitochondria-targeted Fat1 ICD (*Fat1*^{mito}, Extended Data Fig. 5b–d); *Fat1*^{mito} repressed both baseline and maximal OCRs in *Fat1*^{KO} cells (Fig. 3a, Extended Data Fig. 5e, f), demonstrating that the Fat1 ICD can suppress respiration through a mechanism intrinsic to mitochondria. Although expression of the Fat1–IL-2R fusion protein appeared stronger than that of *Fat1*^{mito} (Extended Data Fig. 5a, d), the extent of OCR repression achieved was similar, perhaps indicating that only a subset of the Fat1–IL-2R-derived fragments are active in this capacity. To localize this effect within the respiratory chain, we assessed the activities of isolated respiratory complexes. Complexes I and II, but not IV or V, were more active in the absence of Fat1 (Fig. 3b, c, Extended Data Fig. 6a, b). Moreover, Fat1 ICD species of ~60–100 kDa were observed in association with immunocaptured complex I, II and V, but not complex IV (Fig. 3d, Extended Data Fig. 6c). To see if Fat1 affects respiratory complex integrity, we looked at intact mitochondrial complexes by blue native (BN)–PAGE analysis and found that in *Fat1*^{KO} SMCs, complex I incorporation into supercomplexes is increased, without a significant change in the levels of native complex I (Fig. 3e, f, Extended Data Fig. 6d). It has been proposed that supercomplex conformation facilitates electron transfer without increasing production of reactive oxygen species^{22,23}; however, the factors regulating supercomplex assembly are largely unknown.

Orthogonal resolution of BN gels under denaturing conditions indicated a shift of complex I subunits towards supercomplex conformations in *Fat1*-deficient SMCs and no major differences in native complex II (Extended Data Figs 6e, 7). We observed ~60–100-kDa *Fat1* species co-migrating with complex I, II, III, and V, but not complex IV. Notably, a subset of these species also co-migrated with a supercomplex containing at least complexes I and III (Extended Data Fig. 6e). These findings suggest that Fat1 limits SMC mitochondrial respiration by repressing complex I and II activities, and by opposing the incorporation of complex I into supercomplex structures.

To assess Fat1 relevance to clinical vascular disease, we examined segments of diseased post-mortem human arteries adjacent to sites of prior stent placement. FAT1 expression coincided in many areas with ACTA2, a marker of SMCs (Fig. 4a). As it did in mouse SMCs, FAT1 in cultured human aortic SMCs (HASMCs) co-localized with mitochondria (Fig. 4b). Moreover, knockdown of *FAT1* in HASMCs increased both cell proliferation (Extended Data Fig. 8a–c) and basal and maximal mitochondrial respiration (Extended Data Fig. 8d, e). Thus the effects of Fat1 on SMC mitochondrial activity and growth are conserved across species.

To test SMC Fat1 function *in vivo*, we crossed *Fat1*^{loxP} and *Tagln-cre* mice²⁴ to generate *Fat1*^{loxP/loxP} *Tagln-cre* (denoted hereafter as *Fat1*^{SMKO}, for *Fat1* smooth muscle cell knockout) animals (Extended Data Figs 1b, 9a, b). Baseline arterial structure appeared normal, with minimal Fat1 expression (Fig. 4c, Extended Data Fig. 9a). After vascular injury²⁵, Fat1 appeared in the media and neointima of wild-type arteries, co-localizing with Acta2 and showing a perinuclear distribution consistent with findings in isolated cells (Figs 1d, 4b, Extended Data Fig. 9a, b). Injured arteries from *Fat1*^{SMKO} mice showed markedly higher medial SMC hyperplasia and neointimal growth (Fig. 4c–f), increased phospho-histone H3 and cyclin D1 (Fig. 4g), and increased

staining for nitrotyrosine, a marker of increased superoxide stress (Extended Data Fig. 9c). Moreover, SMCs from injured *Fat1*^{SMKO} vessels exhibited higher OCRs than those from injured control vessels (Fig. 4h, i, Extended Data Fig. 9d). Thus, deletion of Fat 1 in SMCs *in vivo* permits an exaggerated vascular injury response characterized by SMC hyperproliferation and enhanced mitochondrial respiration, which supports the idea that Fat1 induction after vascular injury serves as a physiological regulator of SMC growth, in part by controlling mitochondrial function.

As a large, type-I transmembrane protein, full-length Fat1 appears an unlikely direct regulator of mitochondrial function. Previous studies describe cell-type-dependent cleavage of Fat1 that yields multiple smaller species, including Fat1-ICD-bearing derivatives of ~85 and ~65 kDa²⁶. Although the precise structure of the physiologic mitochondrial Fat1 species that we have identified remains to be determined, *in silico* analysis of potential Fat1 C-terminal peptides that start at amino acids 4195, 4203, or 4478 yield, on the basis of MitoProII²⁷ analysis, respective probabilities of 0.5881, 0.1140, and 0.1548 for export to mitochondria. Moreover, a mitochondrial function for *Drosophila* Fat, an atypical cadherin most closely related to mammalian Fat4 (ref. 28), has recently been reported; a 68 kDa *Drosophila* Fat cytoplasmic fragment interacts with and stabilizes complex I²⁸. This last finding contrasts notably with the increased respiratory activity that we find with Fat1 inactivation, and probably partly reflects the structural divergence between these cadherins²⁹. Nevertheless, the finding that both *Drosophila* Fat and mammalian Fat1 participate in mitochondrial regulation is intriguing and perhaps suggests an ancient shared linkage between primordial cell surface proteins and control of mitochondrial function. Whether Fat4 is active in mitochondria, and how these distinct members of the Fat family might cooperate to regulate mitochondrial function, remain to be determined.

Although the factors or conditions that trigger Fat1 processing and translocation of C-terminal fragments to mitochondria have yet to be defined, our results suggest a relatively direct mechanism for sensing the extracellular milieu and relaying signals to mitochondria to control respiratory chain activity and thereby affect other key metabolic functions. This mechanism could potentially be targeted in hyperproliferative conditions in the cardiovascular system, as shown here, or in other scenarios. To this point, the loss or mutation of Fat1 (refs 8, 9) and the gain of Fat1 expression^{10,30} have both been linked to different cancers. In addition, mutations in the *FAT1* gene have recently been implicated in abnormal renal development¹¹ and skeletal muscle formation¹², and *FAT1* sequence variants have been associated with bipolar disorder, schizophrenia, and autism spectrum disorders¹³. The potential contribution of mitochondrial dysregulation in these divergent scenarios is an area for further study.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 1 September 2015; accepted 19 October 2016.

Published online 9 November 2016.

- Shyh-Chang, N. *et al.* *Lin28* enhances tissue repair by reprogramming cellular metabolism. *Cell* **155**, 778–792 (2013).
- Ahmad, T. *et al.* Miro1 regulates intercellular mitochondrial transport & enhances mesenchymal stem cell rescue efficacy. *EMBO J.* **33**, 994–1010 (2014).
- Bereiter-Hahn, J. Mitochondrial dynamics in aging and disease. *Prog. Mol. Biol. Transl. Sci.* **127**, 93–131 (2014).
- Yu, E. P. & Bennett, M. R. Mitochondrial DNA damage and atherosclerosis. *Trends Endocrinol. Metab.* **25**, 481–487 (2014).
- Ikeda, Y. *et al.* New insights into the role of mitochondrial dynamics and autophagy during oxidative stress and aging in the heart. *Oxid. Med. Cell. Longev.* **2014**, 210934 (2014).
- Pavlidis, S. *et al.* The reverse Warburg effect: aerobic glycolysis in cancer associated fibroblasts and the tumor stroma. *Cell Cycle* **8**, 3984–4001 (2009).
- Martinez-Outschoorn, U., Sotgia, F. & Lisanti, M. P. Tumor microenvironment and metabolic synergy in breast cancers: critical importance of mitochondrial fuels and function. *Semin. Oncol.* **41**, 195–216 (2014).
- Morris, L. G. *et al.* Recurrent somatic mutation of *FAT1* in multiple human cancers leads to aberrant Wnt activation. *Nat. Genet.* **45**, 253–261 (2013).

- Nakaya, K. *et al.* Identification of homozygous deletions of tumor suppressor gene *FAT* in oral cancer using CGH-array. *Oncogene* **26**, 5300–5308 (2007).
- de Bock, C. E. *et al.* The *Fat1* cadherin is overexpressed and an independent prognostic factor for survival in paired diagnosis-relapse samples of precursor B-cell acute lymphoblastic leukemia. *Leukemia* **26**, 918–926 (2012).
- Gee, H. Y. *et al.* *FAT1* mutations cause a glomerulotubular nephropathy. *Nat. Commun.* **7**, 10822 (2016).
- Puppo, F. *et al.* Identification of variants in the 4q35 gene *FAT1* in patients with a facioscapulohumeral dystrophy-like phenotype. *Hum. Mutat.* **36**, 443–453 (2015).
- Cukier, H. N. *et al.* Exome sequencing of extended families with autism reveals genes shared across neurodevelopmental and neuropsychiatric disorders. *Mol. Autism* **5**, 1 (2014).
- Owens, G. K., Kumar, M. S. & Wamhoff, B. R. Molecular regulation of vascular smooth muscle cell differentiation in development and disease. *Physiol. Rev.* **84**, 767–801 (2004).
- Hou, R., Liu, L., Anees, S., Hiroyasu, S. & Sibinga, N. E. The *Fat1* cadherin integrates vascular smooth muscle cell growth and migration signals. *J. Cell Biol.* **173**, 417–429 (2006).
- Miwa, S. *et al.* Low abundance of the matrix arm of complex I in mitochondria predicts longevity in mice. *Nat. Commun.* **5**, 3837 (2014).
- Brand, M. D. & Nicholls, D. G. Assessing mitochondrial dysfunction in cells. *Biochem. J.* **435**, 297–312 (2011).
- Birsoy, K. *et al.* An essential role of the mitochondrial electron transport chain in cell proliferation is to enable aspartate synthesis. *Cell* **162**, 540–551 (2015).
- Sullivan, L. B. *et al.* Supporting aspartate biosynthesis is an essential function of respiration in proliferating cells. *Cell* **162**, 552–563 (2015).
- Perez, J., Hill, B. G., Benavides, G. A., Dranka, B. P. & Darley-Usmar, V. M. Role of cellular bioenergetics in smooth muscle cell proliferation induced by platelet-derived growth factor. *Biochem. J.* **428**, 255–267 (2010).
- Palikaras, K. & Tavernarakis, N. Mitochondrial homeostasis: the interplay between mitophagy and mitochondrial biogenesis. *Exp. Gerontol.* **56**, 182–188 (2014).
- Acin-Perez, R. & Enriquez, J. A. The function of the respiratory supercomplexes: the plasticity model. *Biochim. Biophys. Acta* **1837**, 444–450 (2014).
- Genova, M. L. & Lenaz, G. Functional role of mitochondrial respiratory supercomplexes. *Biochim. Biophys. Acta* **1837**, 427–443 (2014).
- Lepore, J. J. *et al.* High-efficiency somatic mutagenesis in smooth muscle cells and cardiac myocytes in SM22α-Cre transgenic mice. *Genesis* **41**, 179–184 (2005).
- Kumar, A. & Lindner, V. Remodeling with neointima formation in the mouse carotid artery after cessation of blood flow. *Arterioscler. Thromb. Vasc. Biol.* **17**, 2238–2244 (1997).
- Sadeqzadeh, E. *et al.* Dual processing of *FAT1* cadherin protein by human melanoma cells generates distinct protein products. *J. Biol. Chem.* **286**, 28181–28191 (2011).
- Claros, M. G. & Vincens, P. Computational method to predict mitochondrially imported proteins and their targeting sequences. *Eur. J. Biochem.* **241**, 779–786 (1996).
- Sing, A. *et al.* The atypical cadherin fat directly regulates mitochondrial function and metabolic state. *Cell* **158**, 1293–1308 (2014).
- Castillejo-López, C., Arias, W. M. & Baumgartner, S. The fat-like gene of *Drosophila* is the true orthologue of vertebrate fat cadherins and is involved in the formation of tubular organs. *J. Biol. Chem.* **279**, 24034–24043 (2004).
- Dikshit, B. *et al.* *FAT1* acts as an upstream regulator of oncogenic and inflammatory pathways, via PDCD4, in glioma cells. *Oncogene* **32**, 3798–3808 (2013).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank R.N. Kitsis for helpful discussions; A. Jenny for critical reading of the manuscript; X.L. Du for technical help with the Seahorse experiments and for scientific advice; G. Perumal at the Einstein Analytical Imaging Facility for help with electron microscopy imaging; and M.A. Gawinowicz at the Columbia Proteomics laboratory for performing the mass spectrometry analysis. This work was supported by funds from the Diabetes Training and Research Center of Albert Einstein College of Medicine (NIH P60DK20541); funds from the Medical Scientist Training Program (NIH T32-GM007288), Cellular, Molecular Biology, and Genetics Training Grant (NIH T32-GM007491), and an American Medical Association Seed Grant (all to L.L.C.); from the American Heart Association to D.F.R.-B. (pre-doctoral award 11PRE5450002) and to N.E.S.S. (Grant-in-Aid 13GRNT16950064); and from the NIH to L.H. (CA205262) and to N.E.S.S. (HL088104 and HL104518).

Author Contributions L.L.C., D.F.R.-B., P.C., C.M.D., and R.H. generated reagents, performed experiments, and analysed data. B.P.O. performed confocal imaging and analysis of the co-localization studies. V.M. and L.H. imaged and analysed the redox-sensitive ratiometric sensor roGFP. L.G. performed immunohistochemistry on human coronary arteries. M.A.P. and A.F. performed the bioinformatic analysis. L.L.C., D.F.R.-B., and N.E.S.S. designed the study and wrote the paper. L.L.C. and D.F.R.-B. contributed equally to the study. All authors read and approved the final manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to N.E.S.S. (nicholas.sibinga@einstein.yu.edu).

Reviewer Information *Nature* thanks M. Bennett, R. Thorne and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

Protein interaction studies. Tandem affinity purification and mass spectrometry (TAP–MS) was performed using an IL-2R–Fat1-ICD fusion protein¹⁵ (Fat1–IL-2R) cloned into the pCTAP vector (240104, Agilent) to place tandem streptavidin and calmodulin binding peptides (SBP, CBP) at the C terminus of the chimeric protein. A negative control construct included the IL-2R sequence, but not the Fat1 ICD. These constructs were transfected into 293T cells (ATCC), from which total protein lysates were sequentially purified with SBP and CBP beads (InterPlay TAP kit, Agilent). After gentle washing, purified protein complexes were eluted from CBP beads and concentrated (UFC500396, Amicon Ultra), resolved by SDS–PAGE, stained with Coomassie, and submitted for MS/MS analysis (Protein Core Facility, Columbia University) to identify Fat1 ICD associated proteins. MS/MS-validated proteins were scored on the basis of their abundance in the Fat1–IL-2R sample in comparison to the IL-2R control using the ProteoIQ software.

Bioinformatic analysis of the TAP–MS results addressed the 175 potential Fat1 ICD protein interactors with a MASCOT ratio (Fat1 ICD over control) greater than 2. To identify novel interactors, we used information on known protein–protein interactions available in the STRING database³¹. We downloaded all proteins that interacted with the 175 candidate interactors at the lowest confidence level in the database (STRING scores greater than 0.15), and calculated the Page Rank for each node (protein) in the STRING network at different confidence levels (low, 0.15; medium, 0.40; high, 0.70; highest, 0.90) using the Perl module Graph::Centrality::Pagerank, available in the CPAN repository (<http://www.cpan.org/>). Nodes with high PageRank fractions refer to highly connected proteins more likely to be pulled down by chance. From the intersection of proteins with ‘high’ and ‘highest’ STRING confidence scores, we isolated 108 novel Fat1 ICD interactors (Page Rank cutoff of <0.005) and performed enrichment analysis with the DAVID Bioinformatics Resource³². Results were plotted using the EnrichmentMap plugin³³ in Cytoscape 3.1 (ref. 34).

Interaction of the Fat1 ICD and endogenous NDUFS3 was assessed using the IL-2R and Fat1–IL-2R plasmids in TAP-based co-immunoprecipitation and immunoblotting. To evaluate Fat1–prohibitin (PHB) interaction, PHB cDNA amplified by RT–PCR from mouse spleen RNA was cloned into pCS2-6XMyC vector to generate a C-terminally tagged PHB–Myc. DelN Fat1-ICD CMV13–Flag³⁵ and/or PHB pCS2-6XMyC plasmids were transfected into 293T cells. Protein lysates (300 µg) were pre-cleared with Protein G PLUS agarose beads (sc-2002, Santa Cruz) and mouse IgG1 isotype control (M075-3M2, MBL). PHB–Myc was captured with c-Myc beads (sc-40 AC, Santa Cruz), and co-immunoprecipitated proteins were eluted by boiling and assessed for c-Myc and Flag using immunoblotting.

Cell culture and imaging. Wild-type and *Fat1*^{KO} mouse aortic SMCs were isolated from aortas of 5-week-old control and *Fat1*^{SMKO} mice^{11,24} by collagenase–elastase digestion³⁶. SMCs were maintained in Dulbecco’s modified eagle medium (DMEM) containing 20% fetal bovine serum (FBS), 1% L-glutamine, 100 U ml^{−1} penicillin, and 100 µg ml^{−1} streptomycin, and were passaged every three to four days. Cells between passages two and five were used for experiments; these early passage cells were not tested for mycoplasma. To evaluate population growth using the AlamarBlue assay (BUF012A, BIO–RAD), 5,000 cells were plated per well of a 96-well plate and followed for up to 96 h. Absorbances at 570 nm and 600 nm were measured on a Synergy 2 microplate reader (BioTek). Reduction of AlamarBlue correlates with the number of viable cells, and was calculated according to the manufacturer’s specifications and expressed as fold-change with respect to baseline. Where indicated, treatment with 0.02–20 µM rotenone or DMSO (vehicle) was administered at 0 and 24 h. Human aortic smooth muscle cells (HASMCs, Clonetics) were maintained in Medium 199 supplemented as indicated for DMEM, and passaged every four to five days.

To assess proliferation, mouse SMCs were incubated with EdU for six hours before being fixed, permeabilized, and stained for EdU and total DNA (Hoechst stain), as advised by the manufacturer (C10337, ThermoFisher Scientific). Fluorescent signals (excitation:emission (Ex:Em) = 495:519 nm for EdU, 350:461 nm for Hoechst) were measured on a Synergy 2 microplate reader (BioTek), and EdU incorporation was expressed as the ratio of EdU to Hoechst signals. In some studies, wild-type and *Fat1*^{KO} cells were transfected with 25 nM of control or *Ndufs3* small interfering RNA (siRNA) (sense sequence 5′-GCAGAACCGUUUUGAGAUUTT-3′; s86598, Ambion) on 60 mm culture dishes using TransIT-X2 (MIR6000, Mirus). The following day, cells were transferred onto 96-well plates at 2,000 cells per well, and grown for an additional three days before EdU incorporation was measured. HASMCs were transfected with 12 nM of control or *FAT1* siRNA (sense sequence 5′-CAGCGAAGUACAGUACAtt-3′; s5035, Ambion) on 60 mm culture dishes using TransIT-X2 (MIR6000, Mirus). Two days after transfection, cells were transferred onto 96-well plates at 3,000 cells per well and grown for an additional two days before EdU assay.

Co-localization studies were performed as follows: SMCs were seeded onto glass coverslips, stained with 60 nM Mitotracker Red (M-7512, Life Technologies) for 40 min at 37 °C, fixed with 4% paraformaldehyde (PFA) in phosphate buffer saline (PBS) for 15 min at 37 °C, blocked overnight (0.2% milk, 2% normal goat serum, 0.1 M glycine, 2% BSA, and 0.15% Triton-X-100 in PBS), and stained for Fat1 using specific antisera¹⁵ and Alexa Fluor 488 goat anti-rabbit IgG secondary antisera in 0.2% BSA in PBS. SMCs lacking Fat1 and omission of the primary antibody served as controls for signal specificity. Fluorescent signals were visualized using a 4-D spinning-disk confocal microscope (Perkin-Elmer) with a 40× (1.3 NA) objective attached to a digital camera (Orca ER; Hamamatsu). Images were analysed with the Image-J (v1.48v) software.

Electron microscopy of wild-type and *Fat1*^{KO} SMCs was performed in the Albert Einstein College of Medicine Analytical Imaging Facility using standard fixation and imaging procedures. Micrographs were analysed using the Image-J (v1.48v) software.

Mitochondrial studies. Cellular fractions highly enriched for mitochondria were isolated using the Qproteome Mitochondria Isolation Kit (37612, Qiagen). Cells were lysed and further disrupted by passing the lysate through a 26-gauge, blunt-end needle, and mitochondria were isolated using differential centrifugation, followed by density gradient separation. Compartmental markers such as BiP and GAPDH (Fig. 1c, Extended Data Fig. 5c) indicated efficient though not absolute fractionation of mitochondrial components.

Mitochondrial oxygen consumption rate (OCR) was measured using a XF96 Extracellular Flux Analyzer (Seahorse Biosciences) following manufacturer’s instructions. Mouse SMCs were plated on a Seahorse 96-well assay plate at 5,000 cells per well, and tested in minimal assay media containing 5 mM glucose, 4 mM pyruvate, and 2 mM L-glutamine. The following drugs were injected to achieve the indicated final concentration: oligomycin (2 µg ml^{−1}), FCCP (3 µM), and rotenone (2 µM). Total protein was isolated from each well and quantified for normalization after the assay. When indicated, *Fat1*^{KO} SMCs were electroporated (Neon System, Life Technologies) with plasmids encoding Fat1–IL-2R, *Fat1*^{mito}, or empty pcDNA3.1 vector, immediately seeded at 6,000 cells per well onto Seahorse plates, and allowed to recover for three days before measuring OCR. To evaluate respiration in HASMCs, cells were transfected with 12 nM of control or *FAT1* siRNA (s5035, Ambion) on six-well culture plates using TransIT-X2 (MIR6000, Mirus). Three days after transfection, HASMCs were re-plated onto Seahorse plates at 3,000 cells per well, and OCR was measured as described. To evaluate respiration in SMCs after vascular injury, ligated arteries and their uninjured counterparts were collected from control and *Fat1*^{SMKO} mice three days after arterial ligation using the collagenase–elastase method³⁶, seeded on a Seahorse 96-well plate at 15,000 cells per well, and allowed to recover for ~80 h before OCR was measured. OCRs from validated mouse aortic SMC cultures were used to normalize and account for technical variability between experiments.

To target the Fat1 ICD to mitochondria, we assembled the *Fat1*^{mito} construct, which includes the first 58 amino acids (with the mitochondrial targeting sequence) of the mouse thioredoxin-2 protein (GenBank accession NM_019913) and amino acids 4,215–4,591 of mouse Fat1, and cloned it into the pcDNA3.1 expression vector. Mitochondria-specific localization of this fusion protein was confirmed by cellular fractionation (Qiagen) of SMCs and 293T cells, followed by SDS–PAGE and immunoblotting with Fat1 antisera¹⁵.

ATP was measured using a bioluminescent assay (30020-1 Biotium, Inc.) in wild-type and *Fat1*^{KO} SMCs. Micromolar concentrations of ATP were calculated using a standard curve and normalized to percentage of AlamarBlue reduction as a surrogate for cell number to estimate net ATP levels per cell. Data were expressed as fold change respect to wild type.

Reactive oxygen species (ROS) levels were assessed by transfection of wild-type and *Fat1*^{KO} cells with a thiol redox-sensitive ratiometric sensor, reduction-oxidation-sensitive GFP (roGFP, Addgene plasmid 49435)³⁷ using TransIT-X2 (MIR6000, Mirus). After two days, cells were imaged on CELLview glass bottom dishes (5662-7860Q, USA Scientific) containing buffer with divalent cations (125 mM NaCl, 5 mM KCl, 1 mM KH₂PO₄, 5 mM glucose, 10 mM NaHCO₃, 1 mM MgCl₂, 1 mM CaCl₂, and 20 mM HEPES) at 37 °C³⁸. Live cell images were acquired using IX81ZDC wide-field microscope (Olympus) and a single CoolsnapHQ2 (Photometrics) cooled CCD-camera on the bottom 100% throughput port of the microscope. For ratiometric determination of cellular ROS content, cells were sequentially imaged with a 40× magnification objective lens (Olympus; UIS2 40X DIC N/A 1.3) using excitation light from a 100 W Hg arc lamp and bandpass filter FF390/40 (Semrock) (targeting the 400 nm excitation peak of the oxidized form of roGFP) and bandpass filter ET470/40 (Chroma Technology) (targeting the 484 nm excitation peak of the reduced form of roGFP). Fluorescence emissions were detected through a bandpass filter ET525/50 (Chroma Technology). Neutral density filters of 10% to 25% transmission (ND 1.0–0.6) were routinely used to reduce the

intensity of the excitation light to prolong cell viability. The camera exposure times for 390 nm and 470 nm excitations and the respective emission detections at 525 nm were maintained at a 2:1 ratio to maximize the light collection within the dynamic range of the detector. Typically, this was on the order of 200–600 ms, depending on the expression levels of the roGFP probe in a cell. Cells were imaged at 10-s intervals for 5 min, and for an additional 15 min after stimulation with 50 μ M H₂O₂. Images were processed as previously described³⁹. In brief, images were corrected for camera noise and for uneven illumination within the field of view, followed by background subtraction and a histogram-threshold-based masking. At each time point, processed images for each excitation wavelength were divided to obtain the ratiometric (390:470 nm) value, which was then expressed as fold increase above the average baseline (before H₂O₂ stimulation) ratio.

Aspartate levels were measured using a colourimetric assay (ab102512 Abcam) in total cell lysates from wild-type and *Fat1*^{KO} SMCs. Data was corrected by subtracting the background as suggested by the manufacturer. Nanomoles of aspartate per sample were calculated using a standard curve and normalized to mg of protein per sample as a surrogate for cell number, and data were expressed as fold change respect to wild type.

Respiratory complex activities and *Fat1*-complex association were assessed as follows. Mitochondrial complex I activity was evaluated in wild-type and *Fat1*^{KO} SMCs using the complex I activity microplate assay kit (ab109721, Abcam) in accordance with the manufacturer's instructions. The rate of oxidation of NADH to NAD⁺ was measured after immunocapturing complex I from 200 μ g of total protein lysate. Enzymatic (diaphorase) activity was calculated as the change in absorbance (1,000 \times (optical density at 450 nm)) per minute (abbreviated as mOD/min). Complex II activity was evaluated using the corresponding assay (ab109908, Abcam), measuring ubiquinol production after immunocapturing complex II from 100 μ g of total protein lysate, with enzymatic activity measured as a decrease in absorbance at 600 nm per minute (mOD/min). Complex IV activity was tested using the corresponding assay (ab109911, Abcam), with the rate of cytochrome C oxidation measured after immunocapturing complex IV from 70 μ g of total protein lysate, and enzymatic activity measured as a decrease in absorbance at 550 nm per minute (mOD per min). Complex V activity was measured in human aortic SMCs treated with control or *FAT1* siRNA using the ATP synthase enzyme activity microplate assay kit (ab109714, Abcam). Hydrolysis of ATP to ADP and phosphate was measured after immunocapturing complex V from 90 μ g of total protein lysate. Enzymatic activity was measured as a decrease in absorbance at 340 nm per minute (mOD per min). Negative controls were run in every assay, a Synergy 2 microplate reader (BioTek) was used to measure absorbance, and data were expressed as fold-change respect to wild-type cells. After the enzymatic assays, plates were washed with PBS, and proteins were eluted from immunocaptured respiratory complexes by adding 10 μ l per well of RIPA buffer with protease inhibitors plus 6 \times Laemmli protein loading buffer, and incubating the plate at 85 °C for 10 min. Recovered proteins were studied by immunoblotting for *Fat1* fragments and respiratory complex subunits.

Respiratory complex conformations were evaluated by blue native (BN)/SDS-PAGE. First dimension BN-PAGE was performed by solubilizing wild-type and *Fat1*^{KO} mitochondria isolates with 1% Digitonin and complex resolution using NativePAGE 3–12% Bis-Tris gel electrophoresis (Life Technologies). Second dimension analysis involved reducing and alkylating separated proteins, loading the gel strip onto a 10% Tris-Glycine gel, and performing SDS-PAGE, following by immunoblotting for specific complex components and *Fat1*. Native Complex I and supercomplex levels were analysed with the Image-J (v1.48v) software.

Animal models. *Tagln-cre* and *Fat1*^{loxP/loxP} mice have been described previously^{11,24}. We crossed *Tagln-cre* (129SV background) and *Fat1*^{loxP/loxP} (129SV/C57BL6 mixed background) mice to generate *Fat1*^{loxP/loxP} (control) and *Fat1*^{loxP/loxP} *Tagln-cre* (*Fat1*^{SMKO}) mice, using a breeding strategy to produce both genotypes within the same litter. Eight to ten week old male mice were studied. All animals were housed in pathogen-free conditions. Procedures followed rules and regulations of the AAALAC, and were approved by the Institutional Animal Care and Use Committee (IACUC) of Albert Einstein College of Medicine. Vascular injury in mice was performed using the carotid artery ligation model, as described previously²⁵. Mice were anaesthetized with ketamine/xylazine. The left common carotid artery was ligated with 6-0 suture just proximal to the carotid bifurcation, and the unmanipulated right carotid artery served as a control. Carotid arteries were collected at different time points up to 14 days after ligation. For collection, mice were euthanized with ketamine/xylazine and exsanguination, and the central vasculature was flushed with PBS and perfused with 4% PFA in PBS for 7 min. Arteries were removed, post-fixed with 4% PFA in PBS overnight, processed before being embedded in paraffin, and cross-sectioned (5 μ m thickness) for subsequent studies. Morphometric analysis of arteries was performed using the ImageJ software as follows: the area of the lumen, the area inside the internal elastic lamina, and the area inside the external elastic lamina were measured in

pixels, and the areas of the intima and media were calculated. The size of the neointima was expressed as the intima:media ratio. Areas of medial hyperplasia were measured in pixels and expressed as percentage of the total medial area. For immunohistochemical evaluation, arterial sections were deparaffinized and rehydrated. Endogenous peroxidase activity was neutralized before tissues were boiled in antigen retrieval solution (H-3300, Vector Labs). Sections were then blocked (2% BSA, 10% normal goat serum, and avidin blocking solution (Vector Labs) in PBS), incubated with primary antibody (in 2% BSA, 10% normal goat serum, and biotin blocking solution (Vector Labs) in PBS), incubated with biotinylated secondary antibody (Vector Labs) at a concentration of 1:500, incubated with ABC reagent (PK-6100, Vector Labs), incubated with DAB/substrate/chromagen system (K3467, Dako), and finally, counterstained with haematoxylin (H-3404, Vector Labs). Omission of primary antibodies served as controls for signal specificity, and the samples were imaged using a COOLSCOPE digital microscope (Nikon). For immunofluorescence studies, arterial sections were deparaffinized, rehydrated, and boiled in antigen retrieval solution (H-3300, Vector Labs). Sections were then incubated in blocking solution (0.3% Triton-X-100, 2% BSA, and 10% normal goat serum in PBS) overnight. Primary antibodies were applied in blocking solution, and fluorochrome-conjugated secondary antibodies were applied in 0.3% Triton-X-100, 2% BSA in PBS. Omission of primary antibodies served as controls for signal specificity. The sections were mounted with Fluoro-Gel II (with DAPI) mounting medium (17985-50, Electron Microscope Sciences) and visualized with an Axio Observer.Z1 (Zeiss) fluorescence microscope.

Human coronary artery studies. Staining of human coronary artery samples was performed using formalin-fixed paraffin-embedded human coronary artery atherosclerotic plaque sections obtained from the CVPPath Institute Sudden Cardiac Death Registry. To evaluate restenotic samples, we chose artery segments adjacent to the site of a bare-metal stent implanted within the preceding 30 days. H&E stain and immunohistochemistry for *FAT1* and *ACTA2* were performed. The immunohistochemical staining was developed by NovaRED kit (Vector Laboratories, Burlingame, CA). The images were captured by Axio Scan. Z1 (Zeiss, Germany) using a 20 \times objective, and figures were prepared on the HALO image analysis platform (Indica Labs, Corrales, NM).

Antibodies. We used antibodies against *Fat1* (ref. 15) (1:10,000 for western blot (WB), 1:2,000 for immunofluorescence (IF), 1:2,000 for immunohistochemistry (IHC)), *BiP* (1:5,000 for WB, 610979, BD Biosciences), *Flag* (1:15,000 for WB, F3165, Sigma), *cyclin D1* (1:1,000 for WB, 1:200 for IHC, 2978S, Cell Signaling), total OXPHOS cocktail (1:1,000 for WB, ab110413, Abcam), *LC3 I/II* (1:1,000 for WB, 2775S, Cell Signaling), complex I *NDUFA9* (1:1,000 for WB, ab14713, Abcam), complex II 70 kDa *Fp* subunit (1:10,000 for WB, 459200, Invitrogen), complex III subunit core 1 (1:2,000 for WB, 459140, Invitrogen), complex IV *MTCO1* (1:2,000 for WB, ab14705, Abcam), complex V *ATP synthase subunit α* (1:1,000 for WB, 459240, Life Technologies), *ACTA2* (1:200 for IHC, M0851 Clone 184, DAKO), phospho-histone H3 (1:150 for IHC, 9701, Cell Signaling), *transgelin* (1:250 for IF, ab14106, Abcam), *calponin 1* (1:200 for IF, ab46794, Abcam), *pecan1* (1:200 for IF, ab28364, Abcam), anti-mouse IgG HRP (1:10,000, 115-036-072, Jackson ImmunoResearch Labs), Alexa Fluor 488 conjugated anti-rabbit (1:250, 111-545-047, Jackson ImmunoResearch), Alexa Fluor 546 conjugated anti-mouse (1:250, A-11003, Molecular Probes). The following antibodies were from Santa Cruz: complex I *NDUFS3* (1:600 for WB, sc374282), *c-Myc* (1:200 for WB, sc789), β -catenin (1:800 for WB, sc7963), *PGC1- α* (1:150 for WB, sc13067), *PGC1- β* (1:100 for WB, sc373771), *Tfam* (1:150 for WB, sc23588), *Mfn1* (1:600 for WB, sc50330), *Mfn2* (1:600 for WB, sc50331), *Acta2* (1:200 for IF, sc32351), nitrotyrosine (1:250 for IF, sc32757), *GAPDH* (1:8,000 for WB, sc25778), anti-rabbit IgG HRP (1:10,000, sc2030), and anti-goat IgG HRP (1:10,000, sc2056).

PCR. RNA was isolated from wild-type and *Fat1*^{KO} SMCs with Trizol (Life Technologies), and cDNA was generated using Superscript III (Life Technologies). Quantitative reverse-transcription PCR (qRT-PCR) was performed using the SYBR Green system (Roche) on a Lightcycler 480 instrument (Roche). The relative abundance of each transcript was calculated using the $\Delta\Delta C_t$ method, with the *Rpl13a* housekeeping gene as reference. We used the following PCR primers.

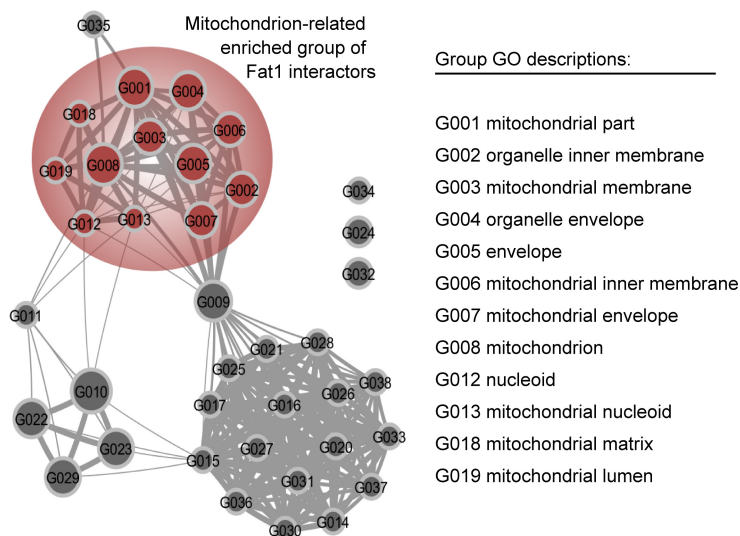
Mouse genotyping. *Fat1*^{loxP} allele forward 5'-GTCTGTGTTTGGCCTG AAGACGTA-3', *Fat1*^{loxP} allele reverse 5'-TTGATGTTGGTGTGGGAAG TGCT-3'; *cre* transgene forward IMR1084 (Jackson Labs), *cre* transgene reverse IMR1085 (Jackson Labs); *Fat1* ^{Δ} allele forward 5'-GTCTGTGTTTGGCCTGAA GACGTA-3', *Fat1* ^{Δ} allele reverse 5'-TAGACAAACCATCACAGGCTGCCT-3'. **qRT-PCR.** *Ppargc1a* forward 5'-GAAAGGGCC AAACAGAGAGA-3', *Ppargc1a* reverse 5'-GTAAATCACACGGCGCTCTT-3'; *Tfam* forward 5'-CAAAGGATGATTCCGGCTCAG-3', *Tfam* reverse 5'-AAGCTGAATATA TGCCTGCTTTTC-3'; *Nrf1* forward 5'-TGGAGT CCAAGATGCTAATGG-3', *Nrf1* reverse 5'-GCGAGGCTGGTTACCACA-3'; *Opa1* forward 5'-ACCAGG

AGAAGTAGACTGTGTCAA-3', *Opa1* reverse 5'-TCTTCAAATAAA CGCAGAGGTG-3'; *Dnm1l* forward 5'-GCTAGTCCACGTTTACCAGA-3', *Dnm1l* reverse 5'-TCCATGTGGCAGGGTCAT-3'; *Mfn1* forward 5'-GTGAG CTTTACCAGTGCAAA-3', *Mfn1* reverse 5'-CACAGTCGAGCA AAAGTAGTGG-3'; *Rpl13a* forward 5'-GCTTACCTGGGCGCTCTG-3', *Rpl13a* reverse 5'-ACATTCTTTTCTGCCTGTTTCC-3'.

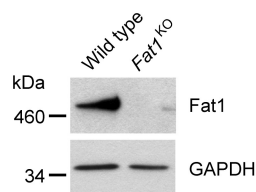
Statistical analysis. In general, at least three independent experiments were performed for each dataset, with a minimum *n* number of three biological replicates. No statistical methods were used to predetermine sample size for *in vitro* experiments. Comparisons between two groups were analysed by two-tailed *t*-test, and comparison between three or more groups were assessed by analysis of variance (ANOVA) followed by post-hoc test where appropriate. Significance was accepted for $P < 0.05$. All statistical analyses were performed and *P* values were obtained using the GraphPad Prism software 6.0. In the animal studies, the sample size was estimated⁴⁰ for a power of 80% and type I error of 0.05 to detect a 1.5- or 1-fold change in the size of neointima measured as the intima:media ratio 3 or 14 days after ligation, respectively. No animals were excluded. The breeding strategy was set up to generate both genotypes of interest within the same litter and mice were allocated into two groups according to their genotypes and not by randomization. The investigator who performed carotid artery ligation remained blinded to group allocation during experiments and outcome assessments.

31. Jensen, L. J. *et al.* STRING 8—a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res.* **37**, D412–D416 (2009).
32. Huang, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protocols* **4**, 44–57 (2009).
33. Merico, D., Isserlin, R., Stueker, O., Emili, A. & Bader, G. D. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS One* **5**, e13984 (2010).
34. Saito, R. *et al.* A travel guide to Cytoscape plugins. *Nat. Methods* **9**, 1069–1076 (2012).
35. Hou, R. & Sibinga, N. E. Atrophin proteins interact with the Fat1 cadherin and regulate migration and orientation in vascular smooth muscle cells. *J. Biol. Chem.* **284**, 6955–6965 (2009).
36. Riascos-Bernal, D. F. *et al.* β -catenin C-terminal signals suppress p53 and are essential for artery formation. *Nat. Commun.* **7**, 12389 (2016).
37. Waypa, G. B. *et al.* Hypoxia triggers subcellular compartmental redox signaling in vascular smooth muscle cells. *Circ. Res.* **106**, 526–535 (2010).
38. Cox, D. *et al.* Requirements for both Rac1 and Cdc42 in membrane ruffling and phagocytosis in leukocytes. *J. Exp. Med.* **186**, 1487–1494 (1997).
39. Spiering, D., Bravo-Cordero, J. J., Moshfegh, Y., Miskolci, V. & Hodgson, L. Quantitative ratiometric imaging of FRET-biosensors in living cells. *Methods Cell Biol.* **114**, 593–609 (2013).
40. Charan, J. & Kantharia, N. D. How to calculate sample size in animal studies? *J. Pharmacol. Pharmacother.* **4**, 303–306 (2013).
41. Procaccio, V. *et al.* Human NDUFS3 gene coding for the 30-kDa subunit of mitochondrial complex I: genomic organization and expression. *Mamm. Genome* **11**, 808–810 (2000).
42. Dieteren, C. E. *et al.* Subunits of mitochondrial complex I exist as part of matrix- and membrane-associated subcomplexes in living cells. *J. Biol. Chem.* **283**, 34753–34761 (2008).
43. Nijtmans, L. G. *et al.* Prohibitins act as a membrane-bound chaperone for the stabilization of mitochondrial proteins. *EMBO J.* **19**, 2444–2451 (2000).
44. Wan, C. *et al.* Panorama of ancient metazoan macromolecular complexes. *Nature* **525**, 339–344 (2015).
45. Steglich, G., Neupert, W. & Langer, T. Prohibitins regulate membrane protein degradation by the m-AAA protease in mitochondria. *Mol. Cell. Biol.* **19**, 3435–3442 (1999).
46. Havugimana, P. C. *et al.* A census of human soluble protein complexes. *Cell* **150**, 1068–1081 (2012).
47. Huttlin, E. L. *et al.* The BioPlex network: a systematic exploration of the human interactome. *Cell* **162**, 425–440 (2015).
48. Guarani, V. *et al.* TIMMDC1/C3orf1 functions as a membrane-embedded mitochondrial complex I assembly factor through association with the MCIA complex. *Mol. Cell. Biol.* **34**, 847–861 (2014).
49. Gilquin, B. *et al.* The AAA⁺ ATPase ATAD3A controls mitochondrial dynamics at the interface of the inner and outer membranes. *Mol. Cell. Biol.* **30**, 1984–1996 (2010).
50. He, J. *et al.* The AAA⁺ protein ATAD3 has displacement loop binding properties and is involved in mitochondrial nucleoid organization. *J. Cell Biol.* **176**, 141–146 (2007).
51. He, J. *et al.* Mitochondrial nucleoid interacting proteins support mitochondrial protein synthesis. *Nucleic Acids Res.* **40**, 6109–6121 (2012).
52. Merle, N. *et al.* ATAD3B is a human embryonic stem cell specific mitochondrial protein, re-expressed in cancer cells, that functions as dominant negative for the ubiquitous ATAD3A. *Mitochondrion* **12**, 441–448 (2012).
53. Christie, D. A. *et al.* Stomatin-like protein 2 binds cardiolipin and regulates mitochondrial biogenesis and function. *Mol. Cell. Biol.* **31**, 3845–3856 (2011).
54. Mitsopoulos, P. *et al.* Stomatin-like protein 2 is required for *in vivo* mitochondrial respiratory chain supercomplex formation and optimal cell function. *Mol. Cell. Biol.* **35**, 1838–1847 (2015).
55. Ewing, R. M. *et al.* Large-scale mapping of human protein–protein interactions by mass spectrometry. *Mol. Syst. Biol.* **3**, 89 (2007).
56. Karblane, K. *et al.* ABCE1 is a highly conserved RNA silencing suppressor. *PLoS One* **10**, e0116702 (2015).
57. Matsuzaki, H., Fujimoto, T., Tanaka, M. & Shirasawa, S. Tespa1 is a novel component of mitochondria-associated endoplasmic reticulum membranes and affects mitochondrial calcium flux. *Biochem. Biophys. Res. Commun.* **433**, 322–326 (2013).
58. Krimmer, T., Rassow, J., Kunau, W. H., Voos, W. & Pfanner, N. Mitochondrial protein import motor: the ATPase domain of matrix Hsp70 is crucial for binding to Tim44, while the peptide binding domain and the carboxy-terminal segment play a stimulatory role. *Mol. Cell. Biol.* **20**, 5879–5887 (2000).
59. Liu, Q., Krzewska, J., Liberek, K. & Craig, E. A. Mitochondrial Hsp70 Ssc1: role in protein folding. *J. Biol. Chem.* **276**, 6112–6118 (2001).
60. Kaul, S. C., Taira, K., Pereira-Smith, O. M. & Wadhwa, R. Mortalin: present and prospective. *Exp. Gerontol.* **37**, 1157–1164 (2002).
61. Hein, M. Y. *et al.* A human interactome in three quantitative dimensions organized by stoichiometries and abundances. *Cell* **163**, 712–723 (2015).
62. Gnoni, G. V., Priore, P., Geelen, M. J. & Siculella, L. The mitochondrial citrate carrier: metabolic role and regulation of its activity and expression. *IUBMB Life* **61**, 987–994 (2009).
63. Chen, Z. & Lash, L. H. Evidence for mitochondrial uptake of glutathione by dicarboxylate and 2-oxoglutarate carriers. *J. Pharmacol. Exp. Ther.* **285**, 608–618 (1998).
64. Palmieri, L. *et al.* Citrin and aralar1 are Ca²⁺-stimulated aspartate/glutamate transporters in mitochondria. *EMBO J.* **20**, 5060–5069 (2001).
65. Fiermonte, G. *et al.* Identification of the mitochondrial glutamate transporter. Bacterial expression, reconstitution, functional characterization, and tissue distribution of two human isoforms. *J. Biol. Chem.* **277**, 19289–19294 (2002).
66. Favre, C., Zhdanov, A., Leahy, M., Papkovsky, D. & O'Connor, R. Mitochondrial pyrimidine nucleotide carrier (PNC1) regulates mitochondrial biogenesis and the invasive phenotype of cancer cells. *Oncogene* **29**, 3964–3976 (2010).

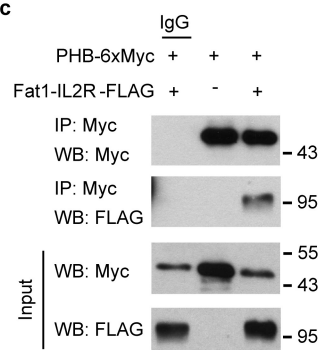
a TAP-MS: String Analysis



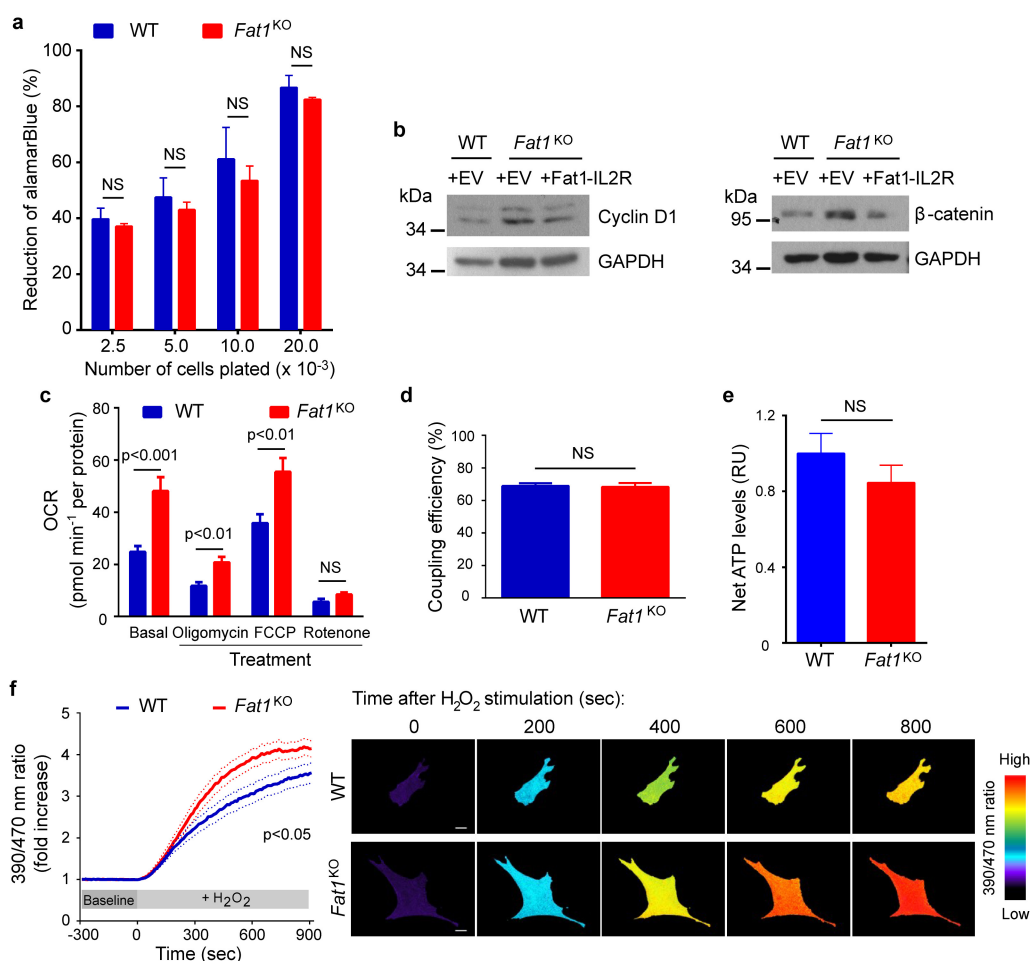
b



c

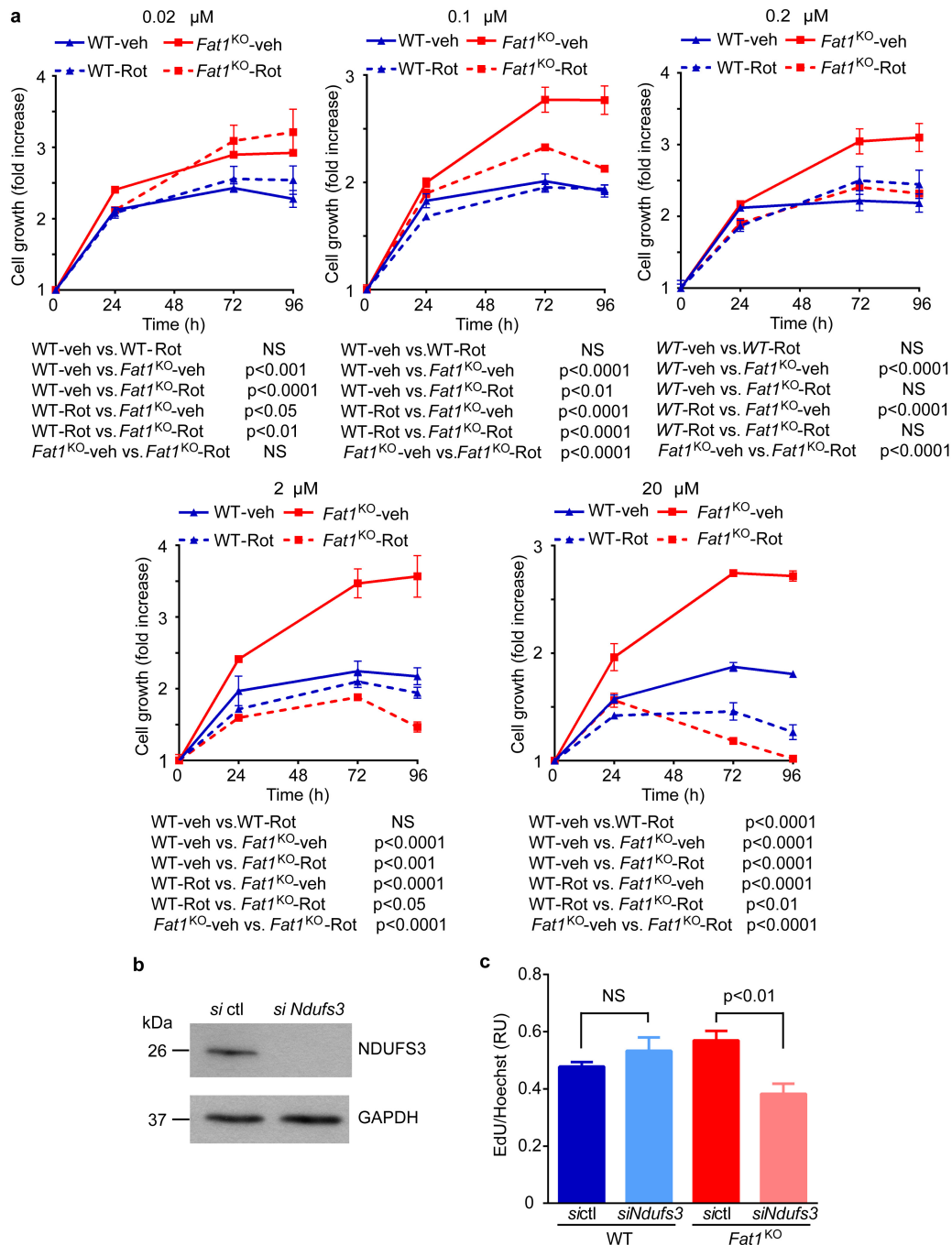


Extended Data Figure 1 | Fat1 ICD associates with inner mitochondrial membrane proteins. **a**, Description of subgroups within the mitochondrial cluster identified by STRING, PageRank, and DAVID enrichment analysis of TAP-MS results. GO, gene ontology. **b**, Western blotting for Fat1 expression in mouse aortic SMCs. **c**, Co-immunoprecipitation of Fat1 ICD and prohibitin (PHB) in 293T cells. IP, immunoprecipitation; performed with IgG (control) or Myc antibody, as indicated. WB, western blot. For gel source data, see Supplementary Fig. 1.



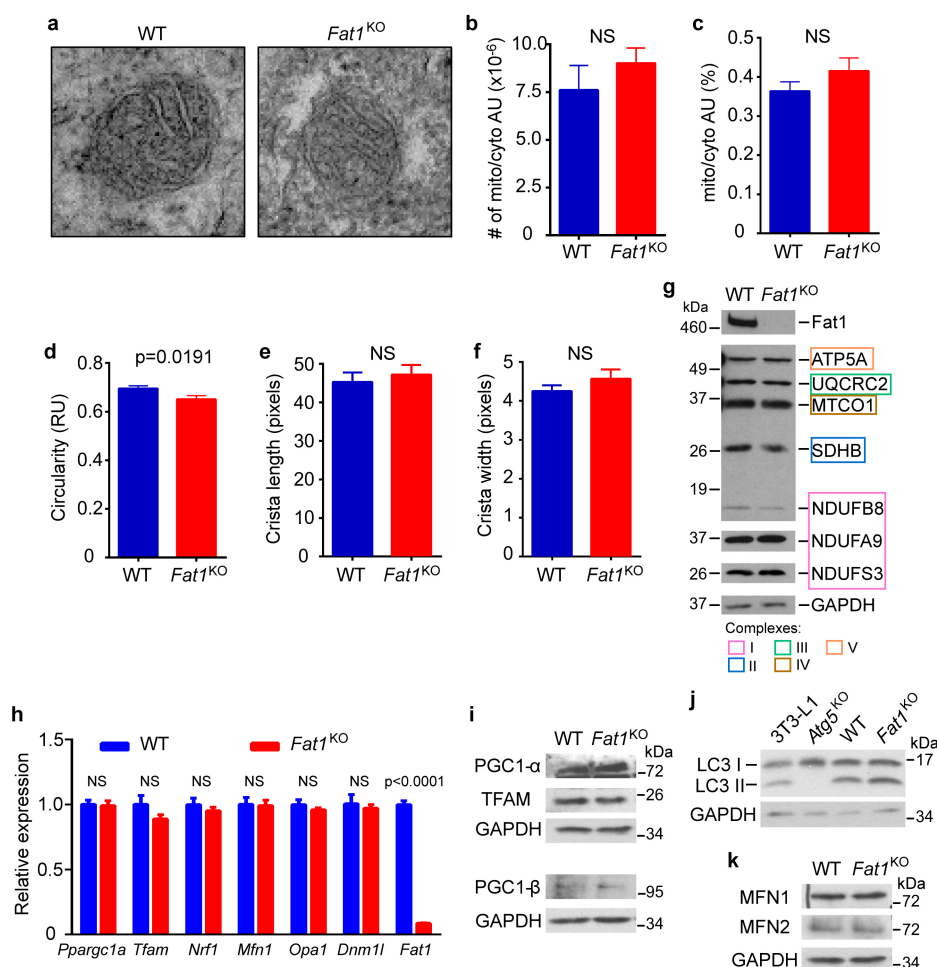
Extended Data Figure 2 | Loss of Fat1 increases expression of pro-proliferative gene products and ATP turnover, but does not affect basal ROS levels in SMCs. **a**, AlamarBlue assay in mouse aortic SMCs plated at specified, matched cell densities, showing correlation of AlamarBlue reduction across a range of input cell numbers ($n = 3$). **b**, Expression of cyclin D1 and β -catenin in mouse aortic SMCs with indicated treatments. EV, empty vector. **c**, Quantification of OCR in mouse aortic SMCs from Fig. 2b ($n = 10$). **d**, Coupling efficiency calculated as percentage of reduction in OCR after oligomycin treatment ($n = 10$). **e**, Net ATP levels in mouse aortic SMCs. RU, relative units ($n = 5$). **f**, ROS levels in mouse

aortic SMCs, measured with a redox-sensitive ratiometric sensor, roGFP. Left, Ratiometric (390/470 nm) value, representing the oxidative state of the sensor and intracellular ROS levels, and expressed as fold increase above the average baseline (before H₂O₂ stimulation) ratio ($n = 15$). Right: Images of cells showing representative ROS levels detected by the ratiometric sensor at baseline ($t = 0$) and at different time points after H₂O₂ stimulation. Data analysed by two-way ANOVA (**a**, **c**, **f**); and two-tailed t -test (**d**, **e**). NS, not significant. All data shown as mean \pm s.e.m. For gel source data, see Supplementary Fig. 1.



Extended Data Figure 3 | *Fat1* suppresses vascular SMC growth by inhibiting the electron transport chain. **a**, Population growth of mouse aortic SMCs in the presence of various concentrations of rotenone, a complex I inhibitor. Addition of rotenone at concentrations from 0.1–2 μM did not compromise wild-type cell growth; by contrast, *Fat1*^{KO} cell growth was suppressed to wild-type levels ($n = 3$); significance assessed by two-way ANOVA. **b**, Western blotting for NDUFS3 expression in

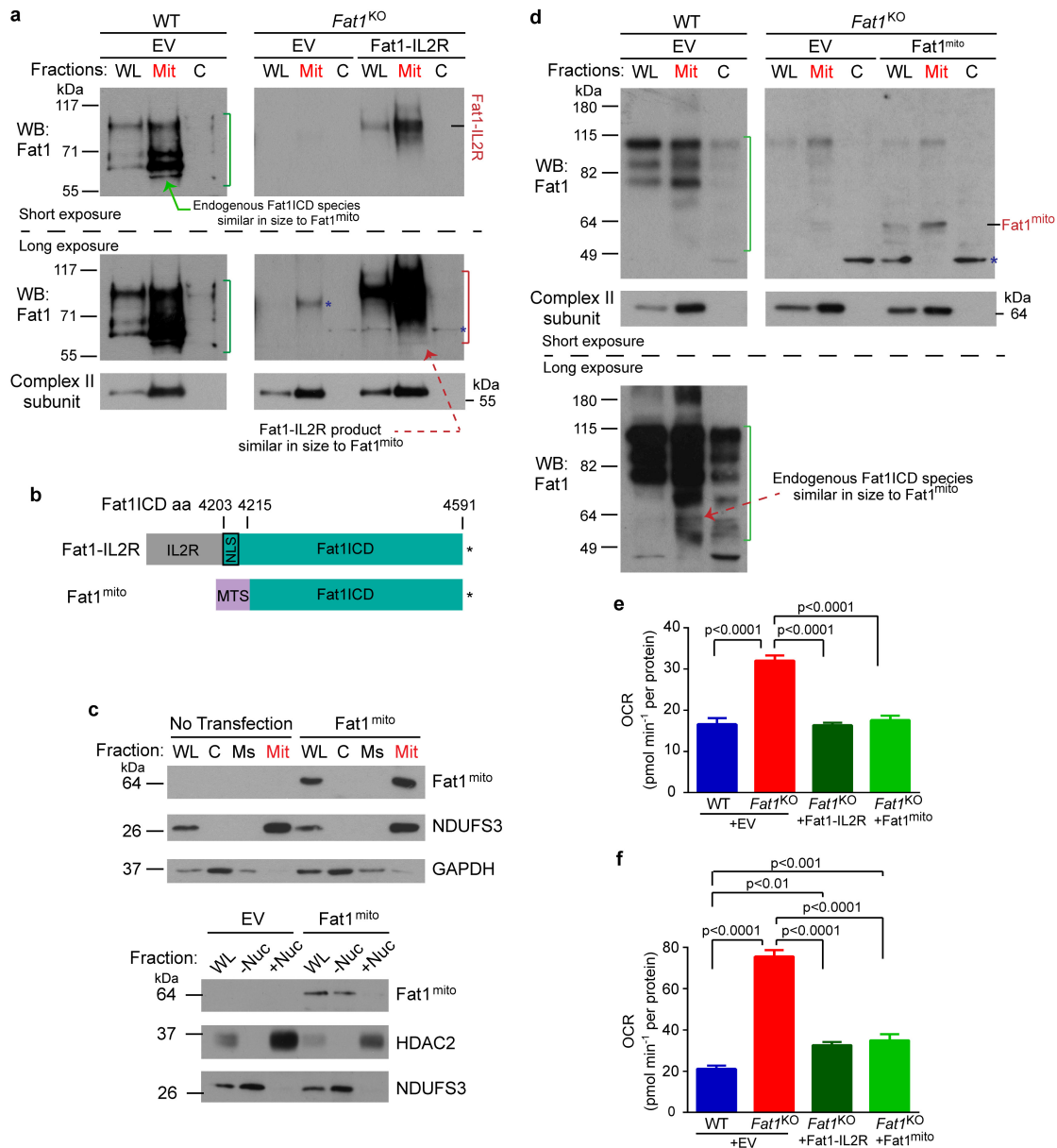
mouse aortic SMCs treated with control siRNA (si ctrl) or *Ndufs3* siRNA (si *Ndufs3*). **c**, Proliferation of mouse aortic SMCs after si *Ndufs3* treatment, expressed as the ratio of EdU to Hoechst signal. RU, relative units. $n = 3$ for *Fat1*^{KO} si *Ndufs3*, $n = 5$ for other groups; significance assessed by one-way ANOVA. NS, not significant. All data shown as mean \pm s.e.m. For gel source data, see Supplementary Fig. 1.



Extended Data Figure 4 | Loss of *Fat1* does not affect overall mitochondrial structure, mass, or dynamics in vascular SMCs.

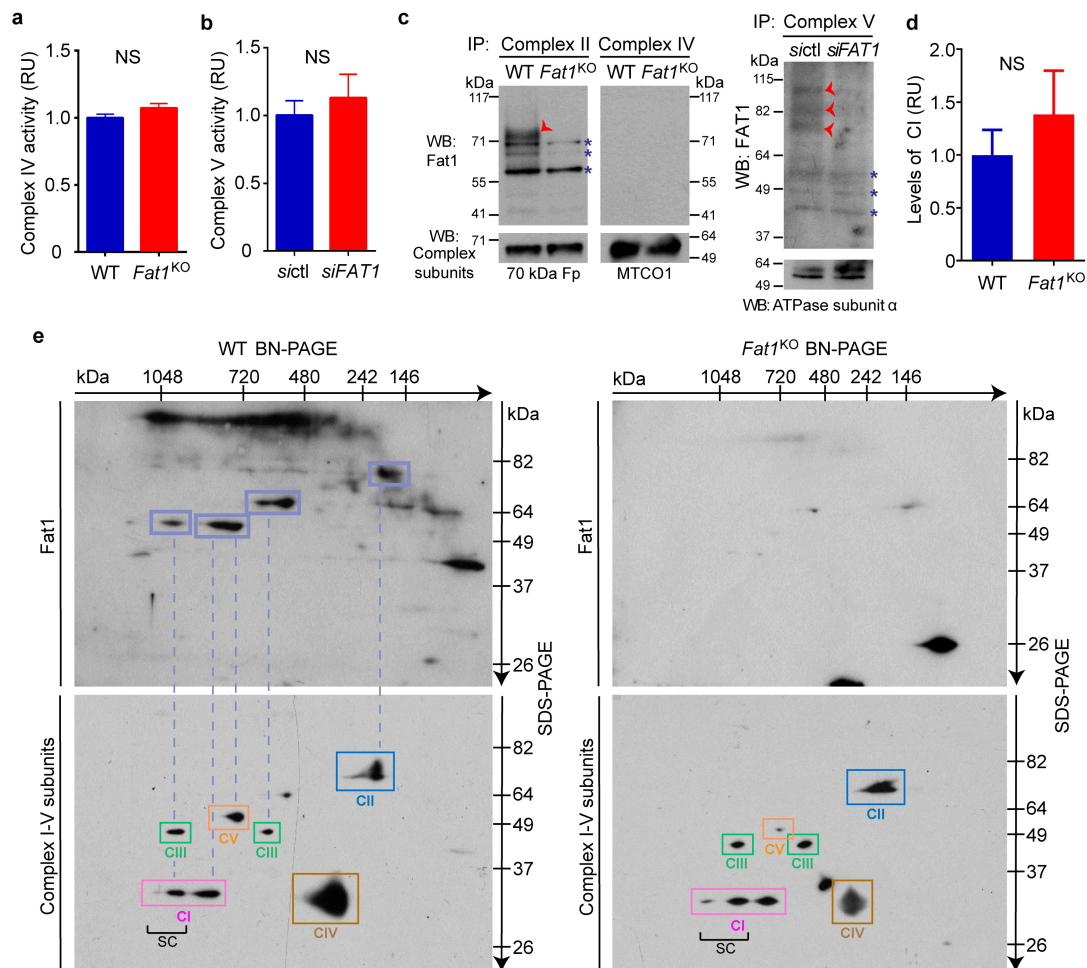
a, Electron microscopy imaging (original magnification, 5,000 \times) of mitochondria in mouse aortic SMCs. **b–f**, Quantification of electron micrographs for the number of mitochondria per cytoplasmic area (**b**, $n=11$); mitochondrial area per cytoplasmic area (**c**, $n=111$ for wild type, $n=119$ for *Fat1*^{KO}); mitochondrial circularity (**d**, $n=111$ for wild type, $n=119$ for *Fat1*^{KO}); mitochondrial crista length (**e**, $n=34$ for wild type, $n=45$ for *Fat1*^{KO}); and crista width (**f**, $n=34$ for wild type, $n=45$ for *Fat1*^{KO}). AU, area units; RU, relative units. **g**, Expression of representative

oxidative phosphorylation proteins from each mitochondrial respiratory complex in total cell lysates of mouse aortic SMCs. **h**, qRT-PCR analysis of mitochondrial biogenesis and fusion/fission markers, normalized to *Rpl13a* expression. $n=3$. **i**, Protein expression of biogenesis markers. **j**, LC3 levels in 3T3-L1 cells and *Atg5* knockout cells (positive and negative controls for autophagy, respectively), as well as in wild-type and *Fat1*^{KO} SMCs. Conversion of LC3 I to LC3 II is indicative of autophagic activity. **k**, Expression of Mitofusins 1 and 2, regulators of mitochondrial fusion. NS, not significant. All data shown as mean \pm s.e.m., significance assessed by two-tailed *t*-tests. For gel source data, see Supplementary Fig. 1.



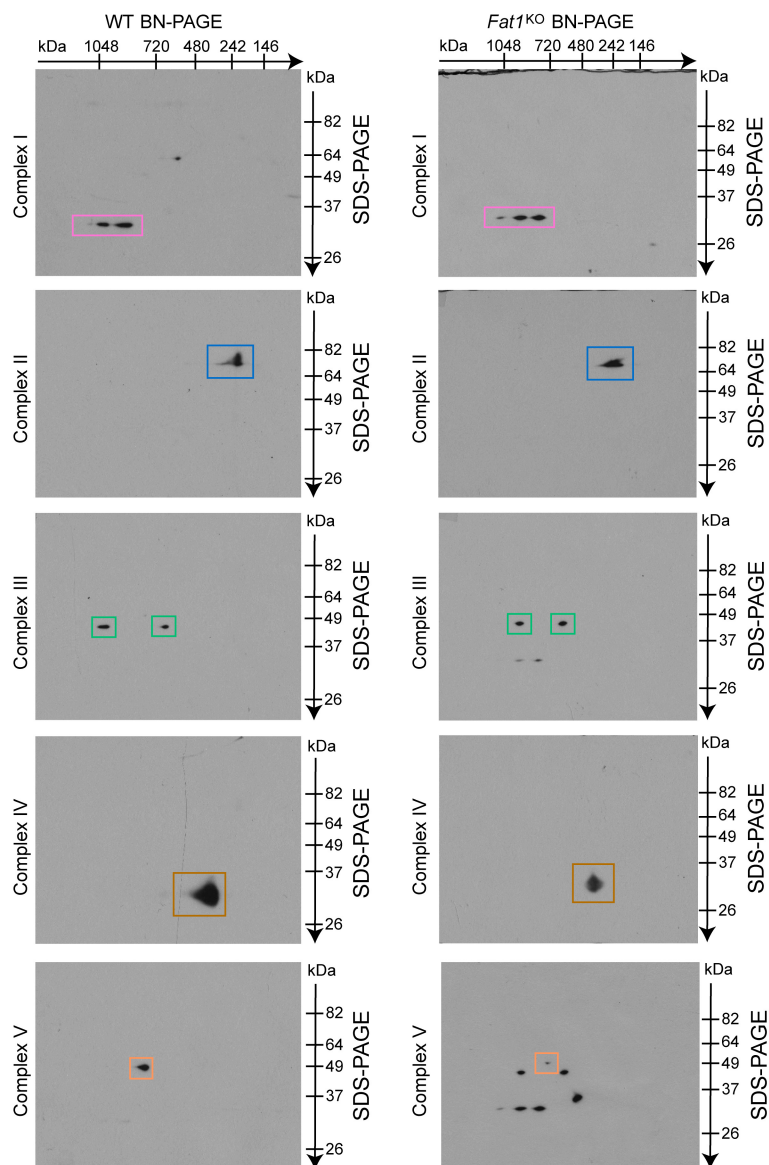
Extended Data Figure 5 | Mitochondria-targeted Fat1 ICD is sufficient to repress oxygen consumption in vascular SMCs. **a**, Electroporation of Fat1-IL-2R in mouse aortic SMCs, followed by subcellular fraction and SDS-PAGE analysis of Fat1 ICD fragments. Green bracket, endogenous Fat1 ICD species; red bracket, Fat1 ICD products from Fat1-IL-2R; blue asterisk, non-specific signal. **b**, Schematic of mitochondria-targeted Fat1 ICD, Fat1^{mito}. Asterisk, stop codon. **c**, Top, Detection of Fat1^{mito} in the mitochondrial fraction of Fat1^{mito}-transfected 293T cells. Bottom, Exclusion of Fat1^{mito} from the nuclear fraction of Fat1^{mito}-transfected

293T cells. **d**, Electroporation of Fat1^{mito} in mouse aortic SMCs, followed by subcellular fraction and SDS-PAGE analysis of Fat1 ICD fragments. Green bracket, endogenous Fat1 ICD species; blue asterisk, non-specific signal. **e**, **f**, Quantification of baseline (**e**) and maximal OCR (**f**) after introducing Fat1-IL-2R or Fat1^{mito} into *Fat1*^{KO} cells from Fig. 3a. Data shown as mean \pm s.e.m., $n = 15$, significance assessed by one-way ANOVA. C, cytoplasmic; EV, empty vector; Mit, mitochondrial; Ms, microsomal; -Nuc, non-nuclear fraction; +Nuc, nuclear fraction; WL, whole-cell lysate. For gel source data, see Supplementary Fig. 1.

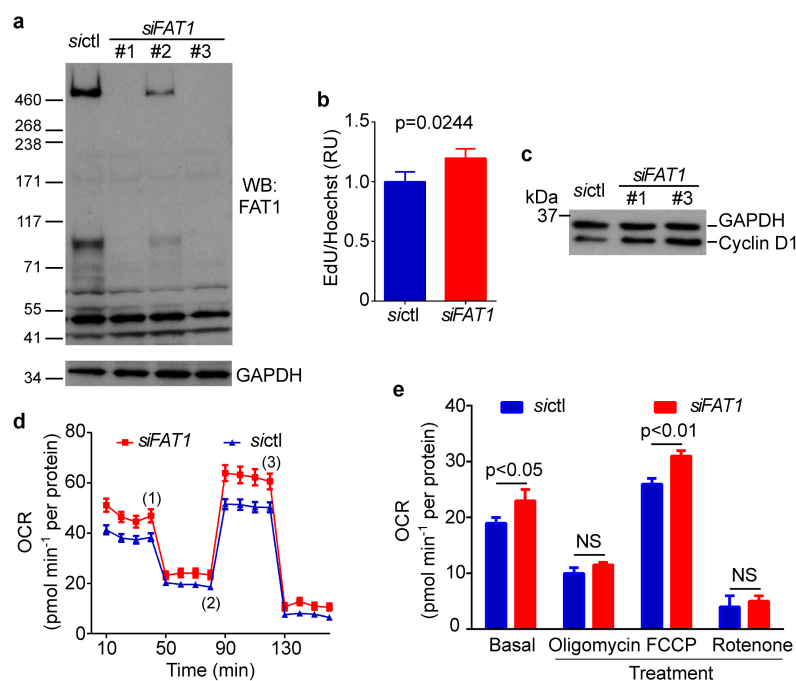


Extended Data Figure 6 | Fat1 ICD associates with mitochondrial respiratory complexes in SMCs, and limits the incorporation of complex I into supercomplexes. **a**, Rate of cytochrome C oxidation by immunocaptured complex IV from mouse aortic SMCs lysates, expressed as fold-change from wild type ($n = 18$). **b**, Rate of hydrolysis of ATP to ADP and phosphate by immunocaptured complex V from human aortic SMCs treated with control siRNA (*siCtrl*) or *FAT1* siRNA (*siFAT1*), expressed as fold change from *siCtrl*. $n = 4$. **c**, SDS-PAGE analysis of proteins eluted from immunocaptured respiratory complexes after completion of enzymatic assays, followed by immunoblotting for Fat1 and complex II, IV, and V subunits. Red arrowhead, specific Fat1 signal;

blue asterisk, non-specific signal. **d**, Quantification of native complex I (CI) levels from BN-PAGE analyses, including the example presented in Fig. 3e. $n = 5$. **e**, Two-dimensional BN/SDS-PAGE analysis of mouse aortic SMC mitochondrial lysates, immunoblotted for Fat1 and complexes I–V. Dashed lines indicate co-migration of Fat1 ICD species with respiratory complexes. Bottom panels are merged images of the individual western blots for complexes I–V (presented in Extended Data Fig. 7). SC, supercomplex. NS, not significant. Data shown as mean \pm s.e.m. and analysed by two-tailed *t*-test (**a**, **b**, **d**). For gel source data, see Supplementary Fig. 1.

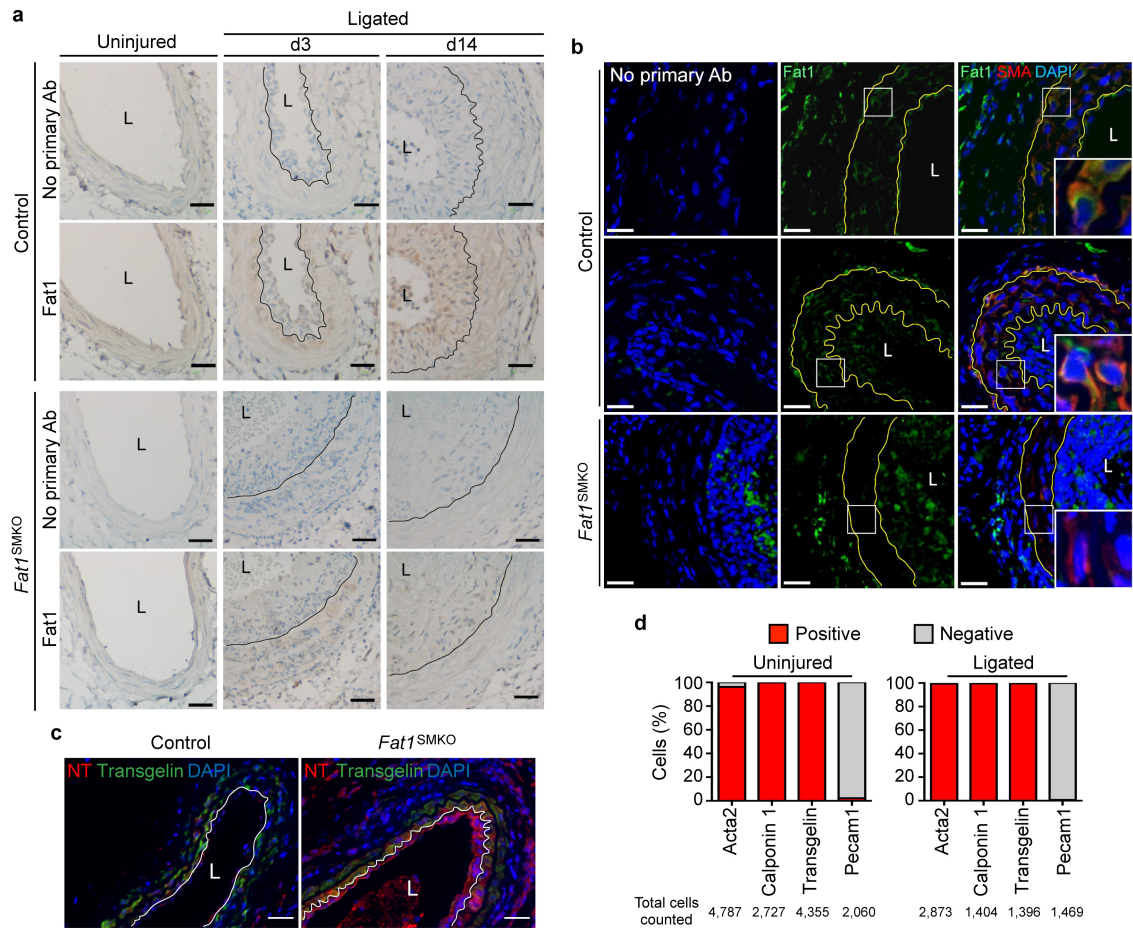


Extended Data Figure 7 | Two-dimensional BN/SDS-PAGE analysis of SMC mitochondrial lysates, immunoblotted for complexes I–V. Individual western blots for complexes I–V used to generate the merged images in Extended Data Fig. 6e, bottom panels.



Extended Data Figure 8 | FAT1 suppresses proliferation and mitochondrial respiration in human SMCs. **a**, Western blotting for FAT1 expression in human aortic SMCs (HASMCs) treated with control siRNA (*sictl*) or *FAT1* siRNAs (*siFAT1*) 1–3. For subsequent experiments, *siFAT1* 3 was used unless otherwise indicated. **b**, Proliferation of HASMCs after *siFAT1* treatment, expressed as the ratio of EdU to Hoechst signal, normalized to *sictl*. $n=3$, significance assessed by two-tailed *t*-test.

c, Expression of cyclin D1 in *sictl*- or *siFAT1*-treated HASMCs. **d**, Oxygen consumption rate (OCR) of *sictl*- or *siFAT1*-treated HASMCs at baseline and in response to $2\mu\text{g ml}^{-1}$ oligomycin (1), $3\mu\text{M}$ FCCP (2), and $2\mu\text{M}$ rotenone (3). **e**, Quantification of OCR from (d). $n=3$, significance assessed by two-way ANOVA. All data shown as mean \pm s.e.m. For gel source data, see Supplementary Fig. 1.



Extended Data Figure 9 | Fat1 expression is induced in SMCs after vascular injury. **a**, Immunohistochemistry for Fat1 in control and *Fat1*^{SMKO} carotid arteries before and after arterial injury, three (d3) or fourteen days (d14) after carotid ligation. L, lumen. The internal elastic lamina has been highlighted with a black line. Scale bar, 25 μ m. **b**, Expression of Fat1 and Acta2 (also known as SMA) by immunofluorescence in control and *Fat1*^{SMKO} carotid arteries, fourteen

days after injury. White squares indicate the regions shown in higher magnification. Yellow line marks the internal and external elastic laminae. Scale bar, 20 μ m. **c**, Nitrotyrosine (NT) staining of control and *Fat1*^{SMKO} carotid arteries, seven days after injury. White line marks the internal elastic lamina. Scale bar, 50 μ m. **d**, Cells isolated from carotid arteries three days after ligation injury expressing SMC (Acta2, calponin 1, transgelin) or endothelial cell (Pecam1) markers by immunofluorescence.

Extended Data Table 1 | Fat1 interactors identified by TAP-MS that localize to the inner mitochondrial membrane

Gene	Protein	Function	Published Interactors	Ref. [†]
<i>NDUFS3</i>	NADH:ubiquinone oxidoreductase core subunit S3	Core subunit of CI that belongs to the minimal assembly required for catalysis.	A number of CI, II, III, and V components	41,42
<i>PHB</i>	Prohibitin	Associates with and stabilizes respiratory complexes, particularly CI. Important for mitochondrial morphology and function, and growth control.	PHB2*, AFG3L2*, TIMM50*, ATAD3A*, UQCRC2*, STOML2*, and many respiratory complex subunits	43,44 45
<i>PHB2</i>	Prohibitin 2	See above	See above	43,44 45
<i>UQCRC2</i>	Ubiquinol-cytochrome c reductase core protein II	Core subunit of CIII, required for its assembly.	NUDFS3* (and other CI subunits), ATP5O* (and other CV subunits), ATP5A1*, PHB*, AFG3L2*, ATAD3A*, HSPD1*	44,46
<i>ATP5F1</i>	ATP synthase, H ⁺ transporting, mitochondrial F0 complex subunit B1	Component of CV. ATP synthesis.	ATP5O*, ATP5A1*, SLC25A3*, STOML2*, HSPD1*, AFG3L2* and many CI, CII, CIII, and CV subunits	44,46 47
<i>ATP5O</i>	ATP synthase, H ⁺ transporting, mitochondrial F1 complex, O subunit	See above.	ATP5F1*, ATP5A1*, UQCRC2*, and other respiratory subunits	44,46
<i>ATP5A1</i>	ATP synthase, H ⁺ transporting, mitochondrial F1 complex, α subunit 1, cardiac muscle	See above.	ATP5F1*, ATP5O*, UQCRC2*, AFG3L2*, STOML2*, and other respiratory subunits	44,46
<i>AFG3L2</i>	AFG3 like matrix AAA peptidase subunit 2	AAA protease involved in IMM proteolysis. Forms a supercomplex with PHB at the IMM. PHB inhibits AFG3L2.	PHB*, ATP5F1*, ATP5A1*, UQCRC2*, SLC25A3*, TIMMDC1* and many CI subunits	44,46 48
<i>ATAD3A</i>	ATPase family, AAA domain containing 3A	Located at contact sites between OMM and IMM. Required for mitochondrial network organization, metabolism, and protein synthesis. Required for nucleoid stability and cell growth. Involved in complex I assembly.	PHB*, UQCRC2*, SLC25A3*, HSPD1*, and other CI and CV subunits.	44,46 49,50 51,52
<i>STOML2</i>	Stomatin like 2	Formation of cardiolipin-enriched microdomains. Recruits PHB to these microdomains to promote PHB function: assembly and stability of respiratory complexes. Promotes supercomplex formation, oxphos, CI and CII activities, and ATP production. Enhances T-cell proliferation.	PHB*, ATP5A1*, ATP5F1*, SLC25A3*, and other CIII and CV subunits	44,46 53,54
<i>TIMMDC1</i>	Translocase of inner mitochondrial membrane domain containing 1	CI assembly: associates with core subunits of CI and components of the CI assembly machinery. Loss of function interrupts CI assembly, supercomplex formation, and mitochondrial respiration.	AFG3L2*, core subunits of CI, CI assembly factors, CIII and CIV subunits	48
<i>TIMM50</i>	Translocase of inner mitochondrial membrane 50	Component of the TIM23 complex; mediates the translocation of transit peptide-containing proteins across the IMM.	PHB*, and some CII subunits	46,55
<i>ABCE1</i>	ATP binding cassette subfamily E member 1	Inhibitor of RNA suppression. Regulation of mRNA turnover. Interacts with various CI subunits and CI assembly factors.		48,56
<i>HSPA9</i>	Heat shock protein family A (Hsp70) member 9	Associates with cristae organizing system complex (MICOS), and modulates mito morphology. In yeast: binds to IMM to regulate mitochondrial import; optimizes translocation acting as an m-AAA protease to degrade/refold misfolded proteins. Implicated in cell proliferation, aging, cell fate determination, differentiation, tumor development.	DNAJA1*, DNAJA2*, DNAJA3*, HSPD1*, CI and CII subunits	44,46 57,58 59,60
<i>NIPSNAP1</i>	Nipsnap homolog 1 (C. elegans)	Molecular function unknown.	Fat1*, HSPD1*	61
<i>SLC25A1</i>	Solute carrier family 25 member 1	Citrate/malate exchange. Important for the bioenergetics of hepatic cells; provides a carbon source for fatty acid and sterol biosynthesis, and NAD ⁺ for the glycolytic pathway.		62
<i>SLC25A3</i>	Solute carrier family 25 member 3	Transport of phosphate groups (along with H ⁺) from the cytosol to the mitochondrial matrix. Regulation of the mitochondrial permeability transition pore.	ATP5F1*, ATAD3A*, AFG3L2*, STOML2*, and many CI, II, III, and V subunits	44,46
<i>SLC25A10</i>	Solute carrier family 25 member 10	Translocation of malonate, malate and succinate in exchange for phosphate, sulfate, sulfite or thiosulfate across IMM.	CI subunits	46
<i>SLC25A11</i>	Solute carrier family 25 member 11	Transport of 2-oxoglutarate across the IMM in exchange for malate or other dicarboxylic acids; important for malate-aspartate shuttle, oxoglutarate/isocitrate shuttle, gluconeogenesis from lactate, and nitrogen metabolism.		63
<i>SLC25A13</i>	Solute carrier family 25 member 13	Calcium-dependent exchange of cytoplasmic glutamate with mitochondrial aspartate across the IMM.		64
<i>SLC25A22</i>	Solute carrier family 25 member 22	Transport of glutamate (with H ⁺) across the IMM.	CII subunits	46,65
<i>SLC25A33</i>	Solute carrier family 25 member 33	Induced by IGF1/mTOR pathway. Controls mtDNA replication and transcription. Promotes mitochondrial respiration and cell growth, and protects from ROS-mediated differentiation during mitochondrial dysfunction.	HSPD1*	44,66

*Associated proteins that were also identified by the Fat1 TAP-MS screen.

[†]Refs 41–66 are cited in this table.

Mechanism of super-assembly of respiratory complexes III and IV

Sara Cogliati^{1*}, Enrique Calvo^{1*}, Marta Loureiro¹, Adela M. Guaras¹, Rocio Nieto-Arellano¹, Carolina Garcia-Poyatos^{1,2}, Iakes Ezkurdia¹, Nadia Mercader^{1,2}, Jesús Vázquez¹ & José Antonio Enriquez^{1,3}

Respiratory chain complexes can super-assemble into quaternary structures called supercomplexes that optimize cellular metabolism¹. The interaction between complexes III (CIII) and IV (CIV) is modulated by supercomplex assembly factor 1 (SCAF1, also known as COX7A2L)². The discovery of SCAF1 represented strong genetic evidence that supercomplexes exist *in vivo*^{2,3}. SCAF1 is present as a long isoform (113 amino acids) or a short isoform (111 amino acids) in different mouse strains^{2,4}. Only the long isoform can induce the super-assembly of CIII and CIV²⁻⁶, but it is not clear whether SCAF1 is required for the formation of the respirasome (a supercomplex of CI, CIII₂ and CIV)^{1,2,4-6}. Here we show, by combining deep proteomics and immunodetection analysis, that SCAF1 is always required for the interaction between CIII and CIV and that the respirasome is absent from most tissues of animals containing the short isoform of SCAF1, with the exception of heart and skeletal muscle. We used directed mutagenesis to characterize SCAF1 regions that interact with CIII and CIV and discovered that this interaction requires the correct orientation of a histidine residue at position 73 that is altered in the short isoform of SCAF1, explaining its inability to interact with CIV. Furthermore, we find that the CIV subunit COX7A2 is replaced by SCAF1 in supercomplexes containing CIII and CIV and by COX7A1 in CIV dimers, and that dimers seem to be more stable when they include COX6A2 rather than the COX6A1 isoform.

We analysed the organization of the respiratory chain by blue native gel electrophoresis (BNGE) in different tissues (Fig. 1a, b). In CD1 mice, CIII and CIV are distributed among several well-characterized structures (IV, IV₂, III₂, supercomplex III₂+IV, supercomplex I+III₂ and supercomplex I+III₂+IV). Other structures of uncertain composition showing co-migration of I and III; I, III and IV, or only IV are also observed. SCAF1 is present in all superassemblies containing CIII and CIV (Fig. 1a). By contrast, C57BL/6J mice showed a depletion of SCAF1 with loss of supercomplex III₂+IV and a concomitant reduction in supercomplex I+III₂+IV, which only remains in relevant amounts in skeletal muscle and heart (Fig. 1b). Further analysis confirmed that supercomplex III₂+IV was virtually absent in C57BL/6J mice while heart, but not liver, retained a significant amount of supercomplex I+III₂+IV (Fig. 1c–e). In adult heart mitochondria from both strains, additional bands containing CIV but not CIII or SCAF1 were detected (Fig. 1f–h). These supernumerary CIV bands were also observed in neonatal C57BL/6J but not in CD1 mitochondria (Fig. 1f–h). The respirasome and supercomplex III₂+IV were present in adult and post-natal day 3 (P3) CD1 hearts while both were absent in P3 C57BL/6J hearts (Fig. 1f–h). Finally, two-dimensional BNGE/SDS-PAGE immunoblotting showed that supercomplex III₂+IV was still undetectable in C57BL/6 heart (Fig. 1i–l) and that the proportion of respirasome was significantly lower than in CD1 heart.

Next, BNGE lanes were sliced into 26 fractions and analysed by mass spectrometry using a data-independent scanning method (DiS)⁷.

We obtained near-complete protein coverage of complex components (Fig. 2a–c). Again, in C57BL/6J mice (SCAF1¹¹¹) supercomplex III₂+IV was virtually absent in C57BL/6J livers and heart while the respirasome was detected in heart mitochondria from both strains (Fig. 2a–c). CIV contains several tissue-specific subunit isoforms; COX7A1 and COX6A2 are heart/skeletal-muscle-specific isoforms whereas COX7A2 and COX6A1 are liver-type isoforms that are present in rodent hearts in significant proportions. SCAF1 is also named COX7A2-like owing to its high similarity to this CIV subunit. We reanalysed the DiS fragmentation data to generate specific quantitative

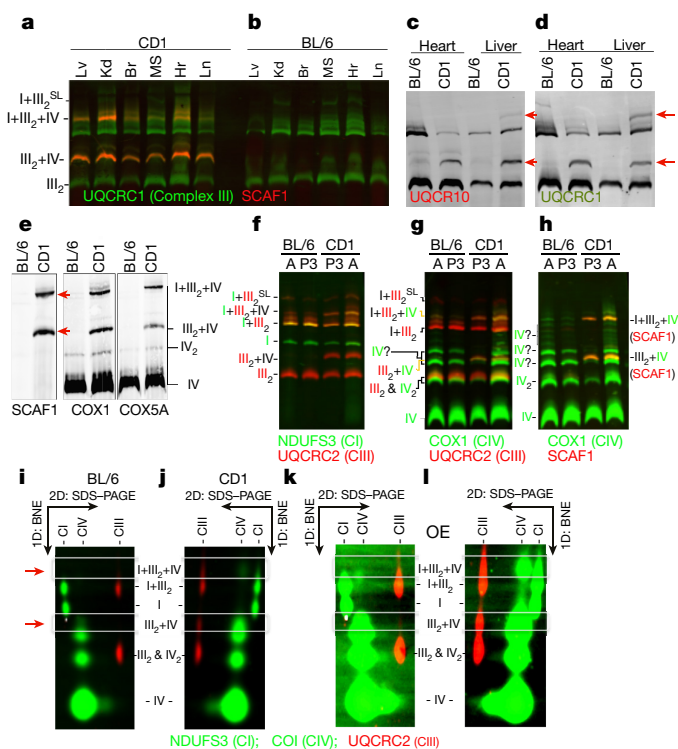


Figure 1 | Super-assembly between CIII and CIV requires functional SCAF1. **a–h**, Immunodetection of the indicated proteins after BNGE of digitonin-solubilized mitochondria from CD1 and C57BL/6J mice (BL/6). Mitochondria were extracted from the indicated tissues (Lv, liver; Kd, kidney; Br, brain; MS, skeletal muscle; Hr, heart; Ln, lung) (**a–d**), from liver (**e**) or from adult (**A**) and 3 days post-natal (**P3**) heart (**f–h**). **i–l**, Immunodetection after BNGE followed by a second dimension SDS-PAGE. OE, overexposed. Red arrows indicate supercomplex III₂+IV and supercomplex I+III₂+IV. Gels representative of 2 (**a, b**), 4 (**c, d**), 2 (**e**) or 1 (**i–l**) biological replicates; in **f–h** representative of 2 biological and 3 technical replicates.

¹Centro Nacional de Investigaciones Cardiovasculares Carlos III, 28029 Madrid, Spain. ²Institute of Anatomy, University of Bern, CH-3000 Bern, Switzerland. ³Departamento de Bioquímica y Biología Molecular y Celular, Universidad de Zaragoza, 50009 Zaragoza, Spain.

*These authors contributed equally to this work.

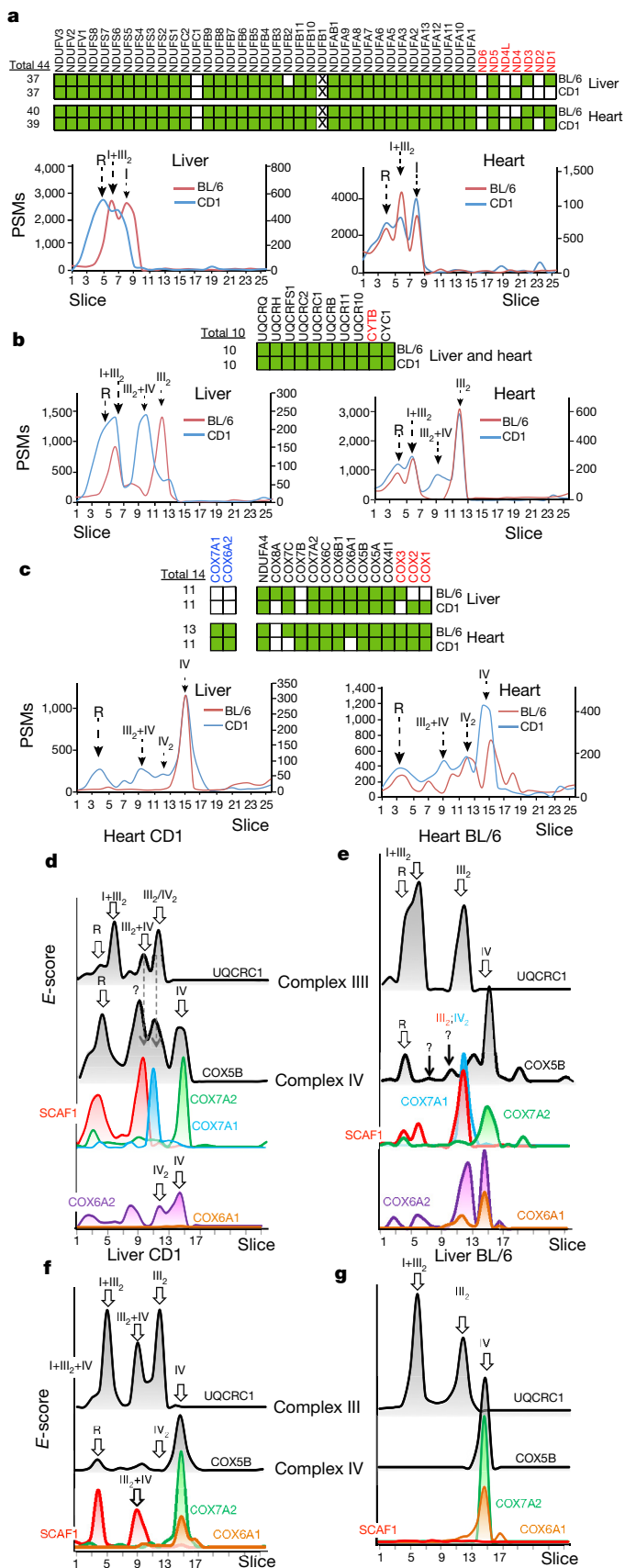


Figure 2 | Super-assembly of the different CIII and CIV structures depends on the subunit composition of CIV. DiS mass spectrometry analysis of BNAGE gel slices of digitonin-solubilized mitochondria from heart and liver from CD1 and C57BL/6J mice. **a–c**, Quantitative profiles of CI (**a**), CIII (**b**) and CIV (**c**), expressed in total peptide-spectrum matches (PSMs); the diagrams indicate the subunits detected of each complex. X boxes indicate proteins not yet described in mice; proteins encoded in nucleus or mitochondria are marked in black or red, respectively; complex IV tissue-specific isoforms are marked in blue; white boxes, not detected; R, respirasome. **d–g**, Quantitative protein profiles, expressed in *E*-scores (see Methods).

in the supercomplex III₂+IV and in the respirasome in both tissues (Fig. 2d, f, red). These results strongly suggest that COX7A2 is replaced by SCAF1 when CIV is assembled into supercomplexes containing CIII.

Notably, this asymmetry in the distribution of the COX7A2 subunit was maintained in C57BL/6J mitochondria, with COX7A2 detected only in the free CIV in both tissues (Fig. 2e, g, green) and slightly in the heart respirasome (Fig. 2e, green). C57BL/6J:SCAF1¹¹¹ was detected in heart and barely in liver (Fig. 2e, g, red), with a pattern different from that of SCAF1¹¹³, co-migrating predominantly with III₂ and supercomplex I+III₂ (Fig. 2e red). These data suggest that, being unable to form stable supercomplex III₂+IV, SCAF1¹¹¹ remained bound to CIII. Recent results proposed that SCAF1 is either a CIII-interacting protein⁶ or a CIV subunit⁸, and our results show that there is no contradiction between these alternatives.

COX7A1 was found only in the dimer form of CIV in CD1 or C57BL/6 hearts, while COX7A2 was found in the monomer (Fig. 2d, e, compare blue and green). Hence, COX7A2 is replaced by COX7A1 to form CIV dimers. Liver mitochondria express only the COX7A2 isoform and the CIV dimer was slightly detectable (Fig. 1e). This suggests that COX7A2 may allow the formation of CIV dimers but with lower efficiency than COX7A1. In support of this interpretation, isolated mice liver mitochondria *in vitro* imported and assembled [³⁵S]COX7A1 but not [³⁵S]COX7A2 into the CIV dimer in addition to the CIV monomer (Extended Data Fig. 5a–d).

We detected COX6A2 in heart mitochondria of both strains while COX6A1 was the isoform detected in liver (Fig. 2d–g, purple and orange). C57BL/6J heart mitochondria expressed both COX6A isoforms (Fig. 2e, purple and orange). COX6A2 was more abundant in CIV dimers and COX6A1 in monomers (Fig. 2e). Notably, when only COX6A1 is expressed (liver mitochondria), the amount of CIV dimer (Fig. 2f, g, yellow trace) is much lower than in heart mitochondria, which also express COX6A2 (Fig. 2d, e). In summary, the CIV dimer seems to be favoured by the concomitant presence of COX7A1 and COX6A2 isoforms.

Our results suggest that SCAF1¹¹¹, the variant carried by C57BL/6J mice, can interact with CIII but not with CIV. COX7A1 and COX7A2 are very similar in size and sequence (Fig. 3a), while SCAF1 is significantly larger, containing in its mature form a long extra domain, as confirmed by proteomics (Fig. 3a and Extended Data Fig. 1). We hypothesized that the region with sequence homology with COX7A2 determines SCAF1 interaction with CIV, while the extra domain is responsible for the interaction with CIII (Fig. 3a), so that SCAF1 acts as bridge between CIII and CIV. To evaluate this hypothesis, we synthesized *in vitro* a series of SCAF1 mutants (Fig. 3b) and studied their incorporation into the different structures in isolated liver mitochondria. [³⁵S]SCAF1¹¹³ was imported into mitochondria and associated with supercomplex III₂+IV and supercomplex I+III₂+IV, but also with supercomplexes I+III₂ and III₂, and marginally with free CIV during short incubation times (Fig. 3c) but not after longer incubation periods (Extended Data Fig. 5). The matrix-exposed region of COX7A harbours His44, a residue critical for its interaction with other subunits of CIV⁹, which corresponds to His73 in SCAF1 (Fig. 3d). Accordingly, the His73Ala SCAF1¹¹³ mutant failed to interact with supercomplex III₂+IV and supercomplex I+III₂+IV, but maintained the ability to interact with supercomplex I+III₂ and III₂ (Fig. 3c, left

profiles of these isoforms and of representative protein components of the complexes (Fig. 2d–g and Extended Data Figs 1–4). Notably, COX7A2 was found almost exclusively in the monomeric CIV form in CD1 heart and liver (Fig. 2d, f, green), while SCAF1 was found

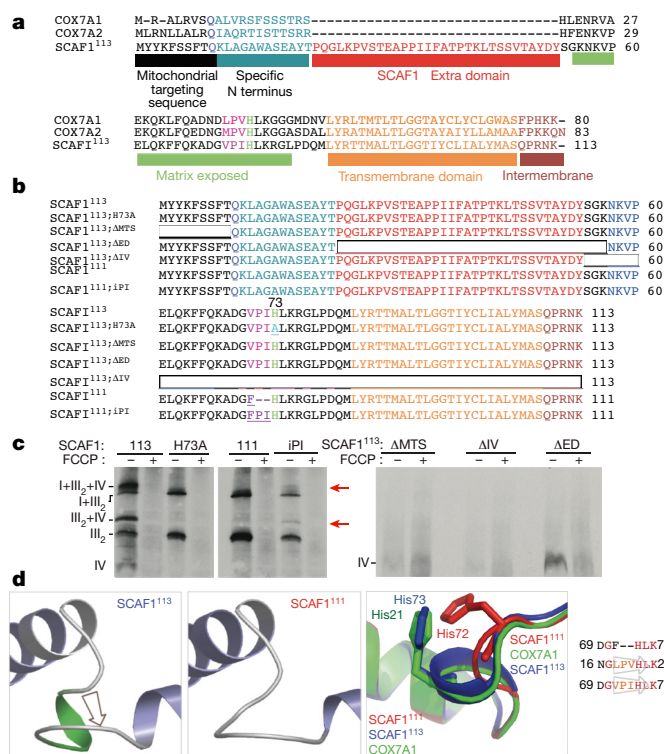


Figure 3 | Functional characterization of SCAF1 domains.

a, b, Alignment of the amino acid sequences of COX7A1, COX7A2 and SCAF1¹¹³ (**a**) and of the mutant constructions (**b**). The critical His residue is highlighted in light green. The surrounding residues, needed for the localization of this His are in purple. Gln residue at the end of the mitochondrial targeting sequence is in blue. The deleted strings of amino acids are boxed and the substitutions or insertions underlined. H73A, His to Ala substitution in SCAF1¹¹³; ΔMTS, deletion of the mitochondrial targeting sequence; ΔED, deletion of the SCAF1 extra domain; ΔIV, deletion of the area similar to COX7A proteins; iPI, Pro and Ile insertion in SCAF1¹¹¹. **c**, Mitochondrial import assay. Autoradiography of BNAGE of CD1 mice liver digitonin-solubilized mitochondria after import of the indicated radiolabelled proteins. Labelling was fully ablated when mitochondrial import was prevented by FCCP-mediated dissipation of the membrane potential. Red arrows indicate supercomplex III₂+IV- and supercomplex I+III₂+IV-containing SCAF1. Representative of 4 biological replicates. **d**, Modelling of the structural consequences of the PI micro-deletion suffered by SCAF1 in C57BL/6J mice (SCAF1¹¹¹), assumed to disrupt the spatial orientation of His73, altering the conformation required for interaction with CIV.

panel). SCAF1¹¹¹ lacks two amino acids adjacent to His73; structural modelling suggested that this deletion strongly affects the relative spatial position of this residue (Fig. 3d). Accordingly, SCAF1¹¹¹ reproduced exactly the phenotype of the His73Ala SCAF1¹¹³ mutant (Fig. 3c, left panel). Reintroduction of two amino acids into the sequence of SCAF1¹¹¹ was sufficient to partially restore the interaction with supercomplex III₂+IV and supercomplex I+III₂+IV (Fig. 3c, middle panel). Elimination of the SCAF1¹¹³ extra domain (SCAF1¹¹³;ΔED) abolished the interaction with the respirasome and supercomplex III₂+IV and also with supercomplexes I+III₂ and III₂ but did not affect the interaction with free CIV (Fig. 3c). Finally, elimination of the region with sequence homology with COX7A2 (SCAF1¹¹³;ΔIV) or the mitochondrial targeting sequence (SCAF1¹¹³;ΔMTS) abolished the import of SCAF1¹¹³ (Fig. 3c, left panel). These results suggest that SCAF1 acts as a bridge between CIII and CIV.

To rule out the possibility that other unknown genetic elements unrelated to SCAF1 could contribute to the observed differences between CD1 and C57BL/6, we backcrossed 129S2/SvPasCrlf (which has SCAF1¹¹³) with C57BL/6J to generate purebred C57BL/6J mice

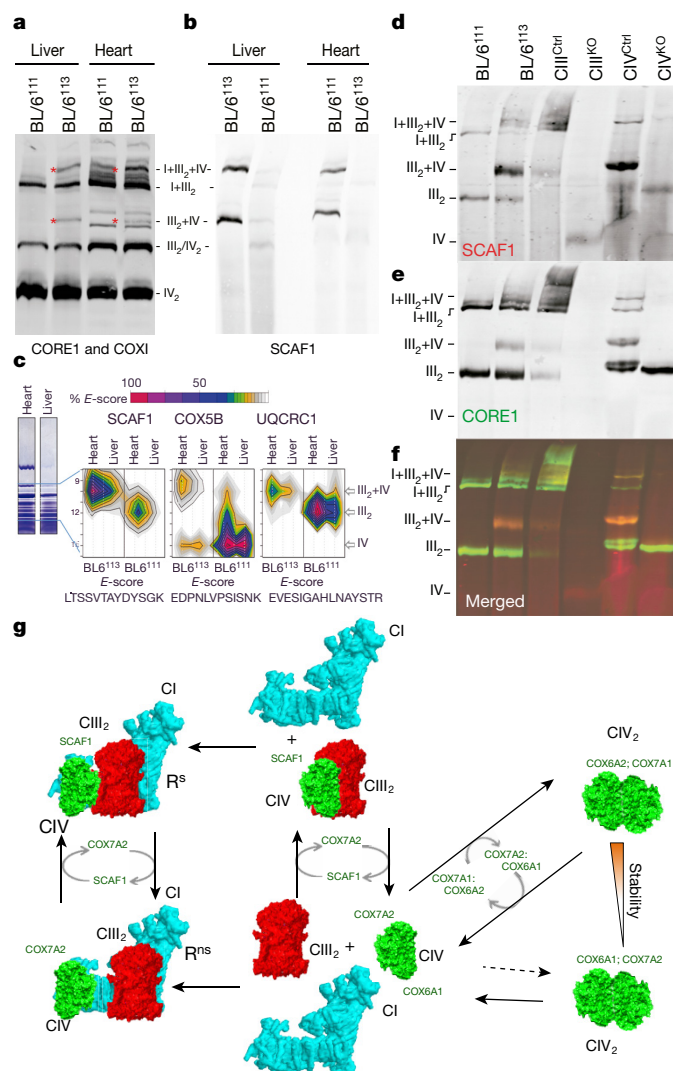


Figure 4 | Recovery of SCAF1 function and models of CIV super-assembly. a–c, Heart and liver digitonin-solubilized mitochondria from C57BL/6J mice harbouring either SCAF1¹¹³ or SCAF1¹¹¹ were subjected to BNAGE and then to immunodetection with the indicated antibodies (**a, b**; representative of 4 biological replicates) or to quantitative DiS mass spectrometry analysis of representative proteins (**c**) in the form of contour plots indicating percentage of maximum E-score of the indicated peptides, normalized according to the level of ATP-synthase within each sample (**d–f**; 1 biological replicates). The same immunodetection analysis was performed with mitochondria from the indicated cell lines. **g**, Proposed model of CIV dimerization and super-assembly driven by the exchange of CIV subunit isoforms.

harbouring SCAF1¹¹³ (B6J.129S2-Cox7a2l). SCAF1¹¹³ was sufficient to recover supercomplex III₂+IV formation in both liver and heart, rescued the respirasome in liver, and increased the respirasome content in heart (Fig. 4a). In these mice, SCAF1¹¹³ was immunodetected in the respirasome and in supercomplex III₂+IV in both liver and heart (Fig. 4b). Proteomics analysis confirmed the assembly of supercomplex III₂+IV containing SCAF1 in both liver and heart of SCAF1¹¹³ C57BL/6J mice (Fig. 4c). Finally, while both supercomplex III₂+IV and the respirasome were detected in C57BL/6-derived fibroblasts containing SCAF1¹¹³, these structures were completely absent in C57BL/6-derived fibroblasts containing SCAF1¹¹¹ (Fig. 4d–f, left lanes). When fibroblasts lacking either CIII or CIV were compared with their respective isogenic control cell lines, we observed that SCAF1 was found to co-migrate with free CIV in the absence of CIII or with CIII in the absence of CIV (Fig. 4d–f). These observations further support

the notion that the co-migration of SCAF1 with structures containing III₂+IV is due to a true interaction with these supercomplexes.

Here we unveil a function for different isoforms of CIV subunits, whose interchange drives the assembly of CIV into different supercomplex structures. In our experimental conditions, COX7A2 and COX6A1 would favour the free CIV form or decrease the stability of the CIV dimer, while COX7A1 and COX6A2 would stabilize the CIV dimer (Fig. 4g). Notably, COX6A2 is located at the interface of the two monomers⁹, where it could directly participate in stabilizing the CIV dimer, while COX7A1 is located away from the interface, so that it would have an indirect conformational role in dimer stabilization (Extended Data Fig. 6a). Similarly, the replacement of COX7A2 by SCAF1 would introduce a bridging motif that allows the binding of CIV to the CIII dimer (Fig. 4g). In fact, computational modelling suggested that this tentative structure is feasible (Extended Data Fig. 6b). Two reports presenting refined cryo-electron microscopy structures of the respirasome have recently been published^{10,11}. One of them speculated that COX7A1 can be substituted for SCAF1 (ref. 10), in full agreement with our experimental conclusions, although the level of resolution achieved did not enable the determination of which of these two proteins was present in the structure. Our results also support the existence of two types of respirasome in our experimental settings, one in which III₂ and CIV are bound through SCAF1 and interact together with CI (SCAF1-mediated respirasome or R^S: I+III₂+IV), and another in which III₂ and CIV interact with CI but not between them (R^{NS}: IV+I+III₂) (Fig. 4g). Notably, two different structures of the respirasome were reported by Letts *et al.*¹⁰ that precisely differed in the presence or absence of an interaction between CIII and CIV. It is tempting to speculate that these two structures may be differentiated by the presence of SCAF1 or COX7A2, determining whether the interaction between CIII and CIV takes place or not. It is likely that R^S would be more stable and abundant than R^{NS}. Our model also reconciles the discrepancies due to SCAF1 microdeletion in C57BL/6J mouse strains with regards to the impact of harbouring SCAF1¹¹¹ or SCAF1¹¹³ in the supercomplexes' organization of the electron transport chain^{2,4}. In addition, our model explains why the predominant CIV-containing supercomplexes have a single copy of CIV, whereas the crystal structure of CIV was obtained from a purified dimer⁹.

In summary, the structural arrangement of the components of the electron transport chain is substantially richer than anticipated and the functional implications of the diversity of complex IV 'flavours' remain as an exciting challenge that awaits further investigation.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 25 May; accepted 17 October 2016.

Published online 24 October 2016.

1. Enríquez, J. A. Supramolecular organization of respiratory complexes. *Annu. Rev. Physiol.* **78**, 533–561 (2016).
2. Lapuente-Brun, E. *et al.* Supercomplex assembly determines electron flux in the mitochondrial electron transport chain. *Science* **340**, 1567–1570 (2013).
3. Ikeda, K., Shiba, S., Horie-Inoue, K., Shimokata, K. & Inoue, S. A stabilizing factor for mitochondrial respiratory supercomplex assembly regulates energy metabolism in muscle. *Nat. Commun.* **4**, 2147 (2013).
4. Mourier, A., Matic, S., Ruzzenente, B., Larsson, N.-G. & Milenkovic, D. The respiratory chain supercomplex organization is independent of COX7a2l isoforms. *Cell Metab.* **20**, 1069–1075 (2014).
5. Williams, E. G. *et al.* Systems proteomics of liver mitochondria function. *Science* **352**, aad0189 (2016).
6. Pérez-Pérez, R. *et al.* COX7A2L is a mitochondrial complex III binding protein that stabilizes the III₂+IV supercomplex without affecting respirasome formation. *Cell Reports* **16**, 2387–2398 (2016).
7. Guarás, A. *et al.* The CoQH2/CoQ ratio serves as a sensor of respiratory chain efficiency. *Cell Reports* **15**, 197–209 (2016).
8. Zhang, K. *et al.* COX7AR is a stress-inducible mitochondrial COX subunit that promotes breast cancer malignancy. *Sci. Rep.* **6**, 31742 (2016).
9. Tsukihara, T. *et al.* The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8 Å. *Science* **272**, 1136–1144 (1996).
10. Letts, J. A., Fiedorczuk, K. & Sazanov, L. A. The architecture of respiratory supercomplexes. *Nature* **537**, 644–648 (2016).
11. Gu, J. *et al.* The architecture of the mammalian respirasome. *Nature* **537**, 639–643 (2016).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported by MINECO: SAF2015-65633-R; BIO2015-67580-P; PRB2 (IPT13/0001-ISCI-SGEFI/FEDER, ProteoRed); RD 12/0042/0054 and RD12/0042/0056), Fundación La Marato TV3 and ERC-Starting Grant (2013 337703 zebraHeart) The CNIC is supported by MINECO and Pro-CNIC Foundation, and is a SO-MINECO (award SEV-2015-0505).

Author Contributions S.C. generated the B6J.129S2-Cox7a2l mice and performed the experiment shown in Figs 1a–d, i, j, 3 and 4a, b, d–f. C.G.-P. and N.M. performed the experiments shown in Fig. 1f–g and Extended Data Fig. 5a–b. R.N.-A. performed the experiments shown in Fig. 1e and cloning of mutants. A.M.G. performed the experiments shown in Extended Data Fig. 5c, d and the preparatory gels for proteomic analysis. E.C. and M.L. performed the proteomic analysis. I.E. performed structural modelling. J.V. and J.A.E. wrote the manuscript and designed the research.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.V. (jvazquez@cnic.es) or J.A.E. (jaenriquez@cnic.es).

Reviewer Information Nature thanks S. Gygi, M. Hüttemann and D. Winge for their contribution to the peer review of this work.

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Molecular biology. The cDNA of SCAF1¹¹³, SCAF1¹¹¹ and the indicated mutants were synthesized and cloned in the pTNT vector by GenScript.

Mitochondrial isolation. Mitochondria were isolated from mouse tissues as indicated elsewhere¹².

Mitochondrial import assay. SCAF1¹¹³, SCAF1¹¹¹, and the different mutant proteins were cloned in the pTNT vector and *in vitro* translated using the TNT T7/SP6 Coupled Reticulocyte Lysate System (Promega) according to the manufacturer's instructions using EXPRE³⁵S Protein Labelling Mix (Perkin Elmer). 3.5 µl of the reaction mix (out of 50 µl) were incubated with 100 µg of isolated liver mitochondria from CD1 mice in 40 µl of reaction buffer (2 mM succinate, 1 mM ATP, 1 mg ml⁻¹ BSA, 1 mM methionine, 40 µg ml⁻¹ chloramphenicol) for 40 min (Fig. 3) at 37 °C. Following this, they were washed two times with 1 ml of reaction buffer to remove the excess of non-imported proteins. Then, the washed mitochondrial pellet was incubated while rotating in 100 µl of reaction buffer for 1 h at room temperature. To increase the incorporation rate and evaluate whether different COX subunits can be incorporated in different forms of supercomplexes, the mitochondria with the radiolabelled proteins were incubated in 500 µl of the same reaction buffer without pre-incubation and during 6 h on rotation at room temperature (Extended Data Fig. 5a–d). In all cases mitochondria were further processed for the BNE and the gel electroblotted onto Hybond-P-polyvinylene fluoride (PVDF) membranes (GE Healthcare) and the radioactive signal detected following exposure of 2–4 days.

Blue native gel electrophoresis (BNGE). Respiratory chain complexes and supercomplexes were analysed by BNGE as described elsewhere¹³.

Western blotting. After electrophoresis, gels were electroblotted onto Hybond-P-polyvinylene fluoride (PVDF) membranes (GE Healthcare) and immunoblotted with specific antibodies against the different subunits of the complexes anti-COX7A2L polyclonal, St John's laboratory (STJ110597); anti-COX7A2L monoclonal, Leitat Technological Center, Biomed Division; anti-COX7A1, Abcam monoclonal (AB123591); anti-COX5B polyclonal, Proteintech (11418-2-AP); anti-COX1, Invitrogen monoclonal (35-8100); anti-UQCRC1, Origene (NM_013387); anti-NDUFS3 monoclonal Abcam (AB14711); anti-UQCRC2 polyclonal Proteintech (14742-1-AC); anti-UQCRC1 Abcam (118687). The secondary antibodies were goat anti-rabbit IgG (H+L), Alexa Fluor 680 conjugate (A-21076) from Life Technologies and anti-mouse dylight 800 from Rockland (610145002) and acquired with the ODYSSEY Infrared Imaging System (LI-COR).

Animal strains. Parental C57BL/6J^{OlaHsd} and CD1 strains were purchased from Harlan Laboratories. Note that the C57BL/6J^{OlaHsd} strain does not harbour the nicotinamide nucleotide transhydrogenase (NNT) spontaneous mutation that renders the encoded enzyme undetectable, and which is characteristic of the C57BL/6J strain provided directly by Jackson Laboratories. B6J.129S2-Cox7a2l (BL6¹¹³) strain was generated by crossing purebred female C57BL/6J mice contained the short allele of COX7A2L (SCAF1¹¹¹) with purebred male 129S2/SvPasCrlf contained the long allele of COX7A2L (SCAF1¹¹³) both in homozygosis. The resulting heterozygous offspring were backcrossed until the F10 to obtain purebred C57BL/6J background mice containing the long form of COX7A2L in homozygous. Offspring after the tenth generation of backcrossing was considered backcrossed. All animal procedures conformed to EU Directive 86/609/EEC and Recommendation 2007/526/EC regarding the protection of animals used for experimental and other scientific purposes, enforced in Spanish law under Real Decreto 1201/2005. Four-month-old male mice were used as a source to purify mitochondria for molecular biology analysis, therefore sample size estimation, randomization and blinding was not necessary.

Genotyping. Total DNA extracted from tails was genotyped by PCR with the following primers: COX7A2L forward CTTTCTTGCTTGCAGAAGG; COX7A2L reverse GAAGGCCTCGTTTCAGGTGG. The product of PCR was resolved in a 15% acrylamide/bisacrylamide gel. The COX7A2L short allele generates a band of 50 bp while the long allele generates a band of 56 bp.

Cell lines. All cell lines were grown in DMEM (D5796 Sigma-Aldrich) supplemented with 5% FBS (Gibco) and 1 mM sodium pyruvate (Lonza) in 5% CO₂, 95% air atmosphere at 37 °C. The different cell lines were generated as described previously¹⁴. The CIII^{KO} cell line and its isogenic control where called A22 and E9 respectively in the original publication¹⁴. The CIV^{KO} cell line and its isogenic control where described elsewhere^{7,15}. BL6¹¹¹ and BL6¹¹³ MAF SV40 transformed cell lines were generated respectively from wild-type C57BL/6J^{OlaHsd} and B6J.129S2-Cox7a2l mice. All cell lines tested negative for mycoplasma.

Mitochondrial proteomics by data-independent scanning mass spectrometry. BNGE gels were excised in 26 slices taking as reference some discrete Coomassie-

stained bands from the same sample: slice 6 corresponds to a band that mainly contains supercomplex I+III₂, slice 10 to CV (ATP synthase), slice 12 to free supercomplex III₂, and slice 15 to free CIV. All slices were in-gel digested with trypsin. The resulting tryptic peptide mixtures were subjected to nano-liquid chromatography coupled to tandem mass spectrometry using a data-independent scanning (DiS) method as described previously⁷. In brief, peptides were injected onto a C-18 reversed phase (RP) nano-column (75 µm I.D. and 50 cm, Acclaim PepMap, Thermo Fisher, San José, CA, USA) and analysed in a continuous acetonitrile gradient consisting of 8–31% B for 130 min, 50–90% B for 1 min (B = 0.5% formic acid in acetonitrile). Peptides were eluted from the RP nano-column at a flow rate of ~200 nl min⁻¹ to an emitter nanospray needle for real-time ionization and peptide fragmentation in either a Q-Exactive or a Q-Exactive HF mass spectrometer (Thermo Fisher). Each sample was analysed in two chromatographic runs covering different mass ranges (from 400 to 750 Da, and from 750 to 1,100 Da, respectively). The DiS cycle consisted of 175 sequential HCD MS/MS fragmentation events with 2-Da windows that covered the whole 350 Da range. HCD fragmentation was performed at a resolution of 17,500 and a maximum injection time of 80 ms with the AGC set to a target of 3 × 10⁵ ions. The whole cycle lasted 30 s or less depending on ion intensity during chromatography. The narrow windows used for fragmentation allowed peptide identification using conventional shotgun searching algorithms. This was done using Sequest running under Proteome Discoverer 1.4 (Thermo Fisher Scientific), allowing two missed cleavages, and using 2 Da and 20 p.p.m. precursor and fragment mass tolerances, respectively. Met oxidation and Cys carbamidomethylation were selected as dynamic or static modifications, respectively. 1% FDR was used as a criterion for peptide identification; this parameter was estimated using a separate inverted database and the refined method¹⁶.

Quantitative proteomics analysis and automated validation of MS/MS spectra.

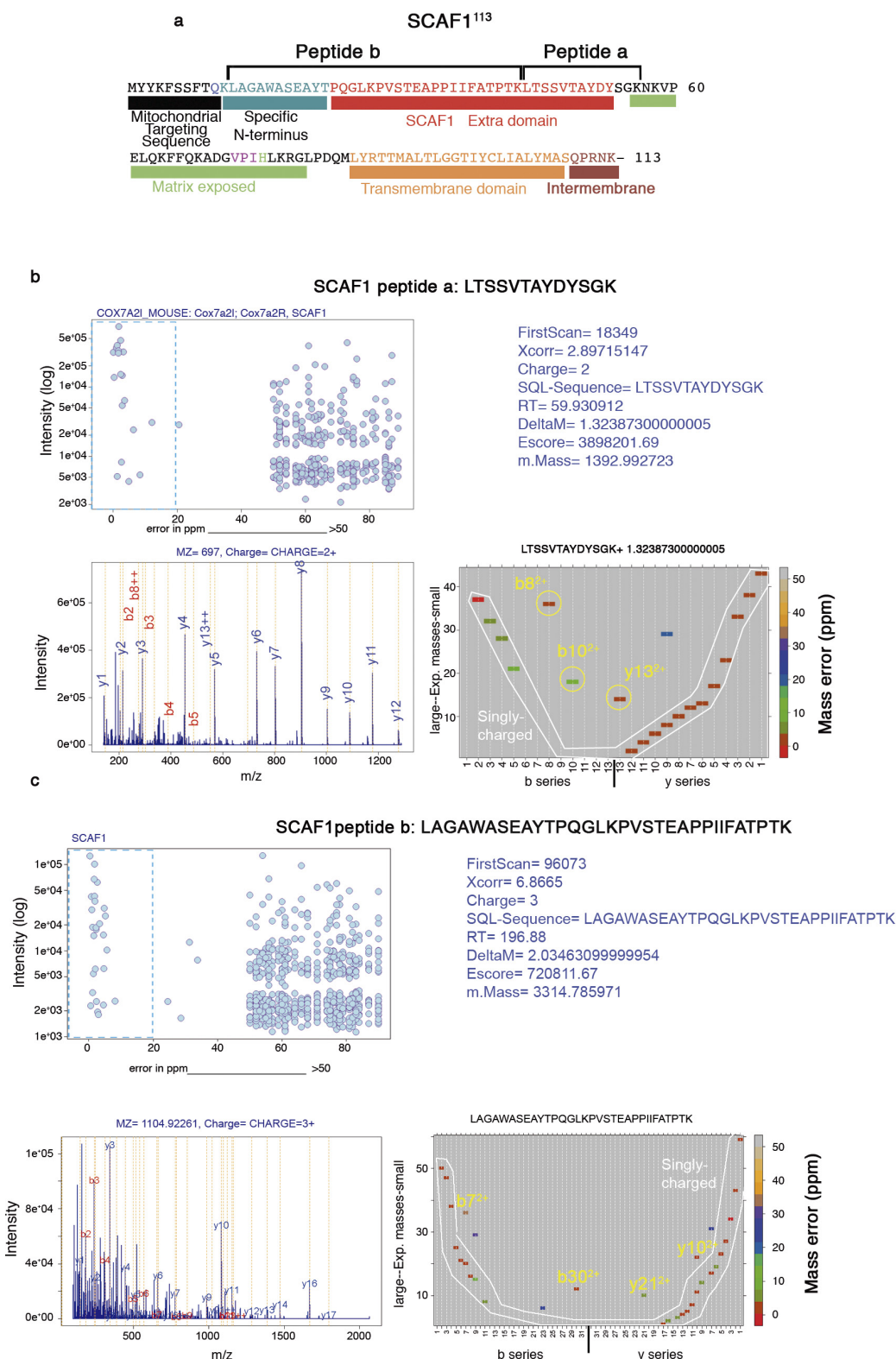
Two different approaches were used. For quantitative analysis of OxPhos complexes we used the total spectral counts¹⁷ of the proteins belonging to each complex. For specific protein quantitation we used the sum of intensities of all the observed fragment ions of charge state from +1 to +3 that were matched within 20 p.p.m. to theoretical fragmentation series of representative proteotypic peptides, which we have called the *E*-score. The sequence of peptides used to draw these traces, together with their MS/MS spectra, showing the specificity of the quantified fragments, are shown in Extended Data Figs 1–4. All *E*-score traces in Fig. 2 are drawn at the same relative intensity scale and are normalized according to the level of ATP synthase within each sample (Extended Data Fig. 5e, f). A custom application in R (Vseq) was developed to calculate the *E*-scores and to generate quality control plots to monitor the specificity of peptide quantitation (see Extended Data Figs 1–4).

Structural modelling. Homology models for COX6A1 and COX7A2 were generated using Modeller¹⁸, on the basis of protein templates from PDB 2Y69 corresponding to chains G and J, respectively. The tentative structural model of SCAF1-mediated supercomplex III₂+IV formation was generated based on homology with COX7A1 (PDB accession code 1V54) and UQCRC1 (PDB accession code 1BGY) using Modeller. The two domains were oriented according to the fitted model for bovine mitochondrial supercomplex I+III₂+IV by single particle cryo-EM on PDB 2YBB, using Pymol and focusing in the III₂+IV adjustment (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC).

Data availability. All the data that support the findings of this study including mass spectrometry data are freely available from the corresponding authors upon request.

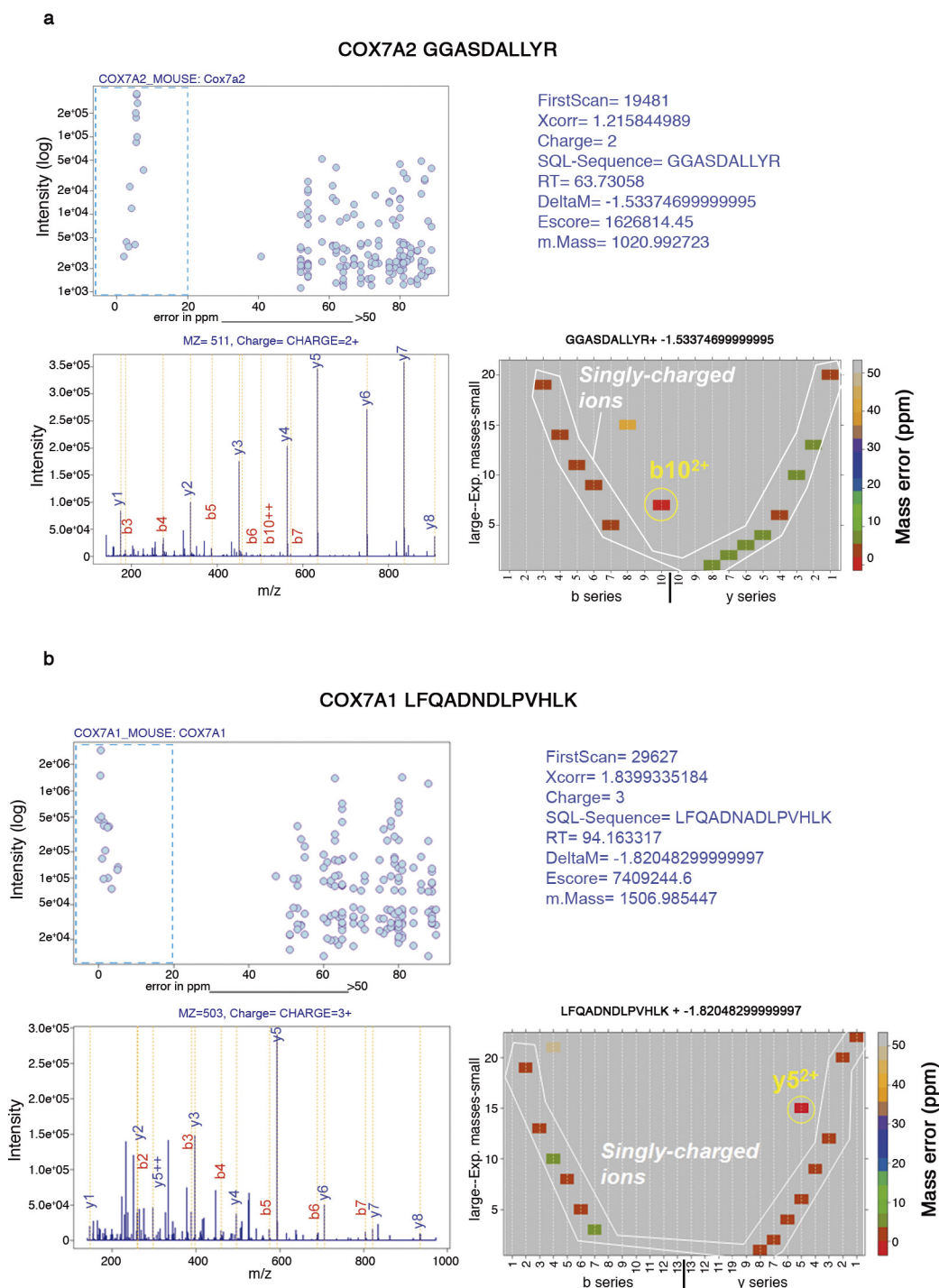
Code availability. Vseq application (R code) is freely available upon request. Source data for the blots/gels in Figs 1, 3 and 4 and Extended Data Fig. 5b, and *E*-score data represented in Fig. 2 are provided in the Supplementary Information.

- Fernández-Vizarra, E. *et al.* Isolation of mitochondria for biogenetical studies: An update. *Mitochondrion* **10**, 253–262 (2010).
- Acín-Pérez, R., Fernández-Silva, P., Peleato, M. L., Pérez-Martos, A. & Enriquez, J. A. Respiratory active mitochondrial supercomplexes. *Mol. Cell* **32**, 529–539 (2008).
- Acín-Pérez, R. *et al.* Respiratory complex III is required to maintain complex I in mammalian mitochondria. *Mol. Cell* **13**, 805–815 (2004).
- Díaz, F., Fukui, H., García, S. & Moraes, C. T. Cytochrome c oxidase is required for the assembly/stability of respiratory complex I in mouse fibroblasts. *Mol. Cell. Biol.* **26**, 4872–4881 (2006).
- Navarro, P. & Vázquez, J. A refined method to calculate false discovery rates for peptide identification using decoy databases. *J. Proteome Res.* **8**, 1792–1796 (2009).
- Nahnsen, S., Bielaw, C., Reinert, K. & Kohlbacher, O. Tools for label-free peptide quantification. *Mol. Cell. Proteomics* **12**, 549–556 (2013).
- Webb, B. & Salí, A. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinformatics* **47**, 5.6.1–32 (2014).



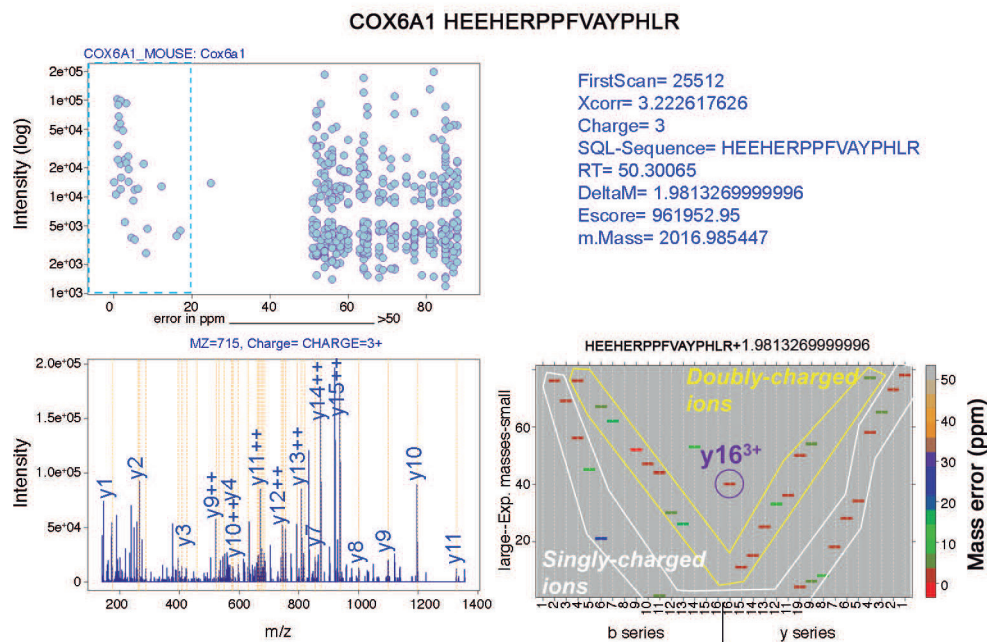
Extended Data Figure 1 | Specificity of tandem mass spectrometry (MS/MS) fragmentation series used to generate quantitative profiles for SCAF1-specific peptides. **a**, Amino acid sequence of SCAF1¹¹³; boxes indicate structural domains. The location of the proteotypic peptides identifying the extra domain (**b**) and the N-terminal region of SCAF1 (**c**) are indicated. **b**, **c**, Quality control plots for proteotypic peptide fragmentation were generated using Vseq, a program written in R. Top, intensity versus fragment mass error plot; only the fragments with a mass error below 20 p.p.m. were assumed to be correct matches. Summation of the intensities located inside the blue dashed box represents the *E*-score.

Lower left, representative MS/MS spectrum of the peptide indicating the matched fragments. Lower right, diagram showing the colour coded mass error of the fragments ranked according their *m/z* values and their correspondence with theoretical fragmentation series. Grey background colour indicates errors higher than 50 p.p.m. The b and the y fragmentation series form a V-shaped path that joins fragments with the same charge. The completeness of the 'V' shape assesses the quality of the identification. Note that since HCD fragmentation was used, y-type fragment ions are more intense than b-type ions.

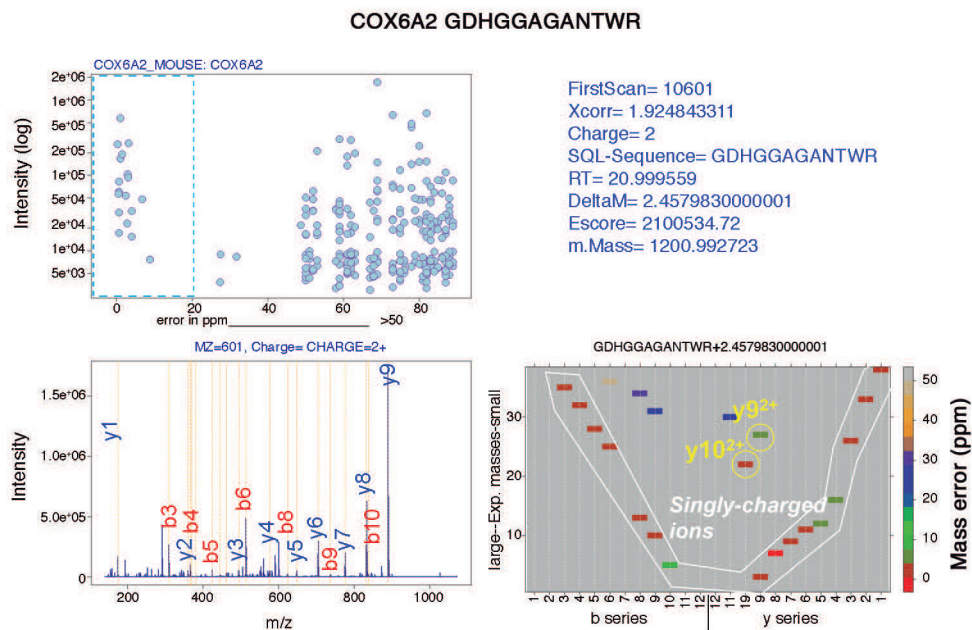


Extended Data Figure 2 | Specificity of MS/MS fragmentation series used to generate quantitative peptide profiles for COX7A2 and COX7A1.
a, b, Profiles for COX7A2 (**a**) and COX7A1 (**b**). The quality control plots are as in Extended Data Fig. 1.

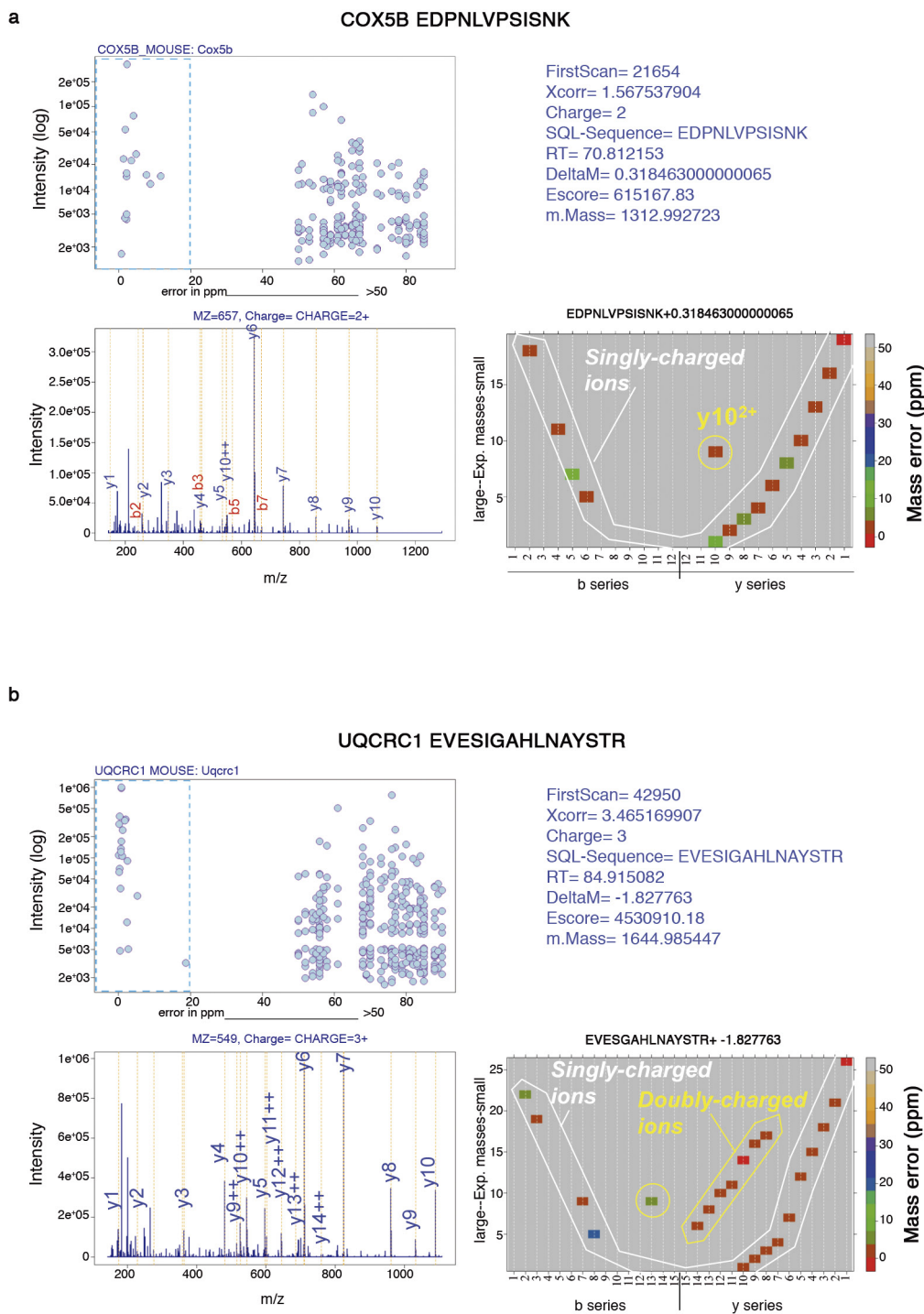
a



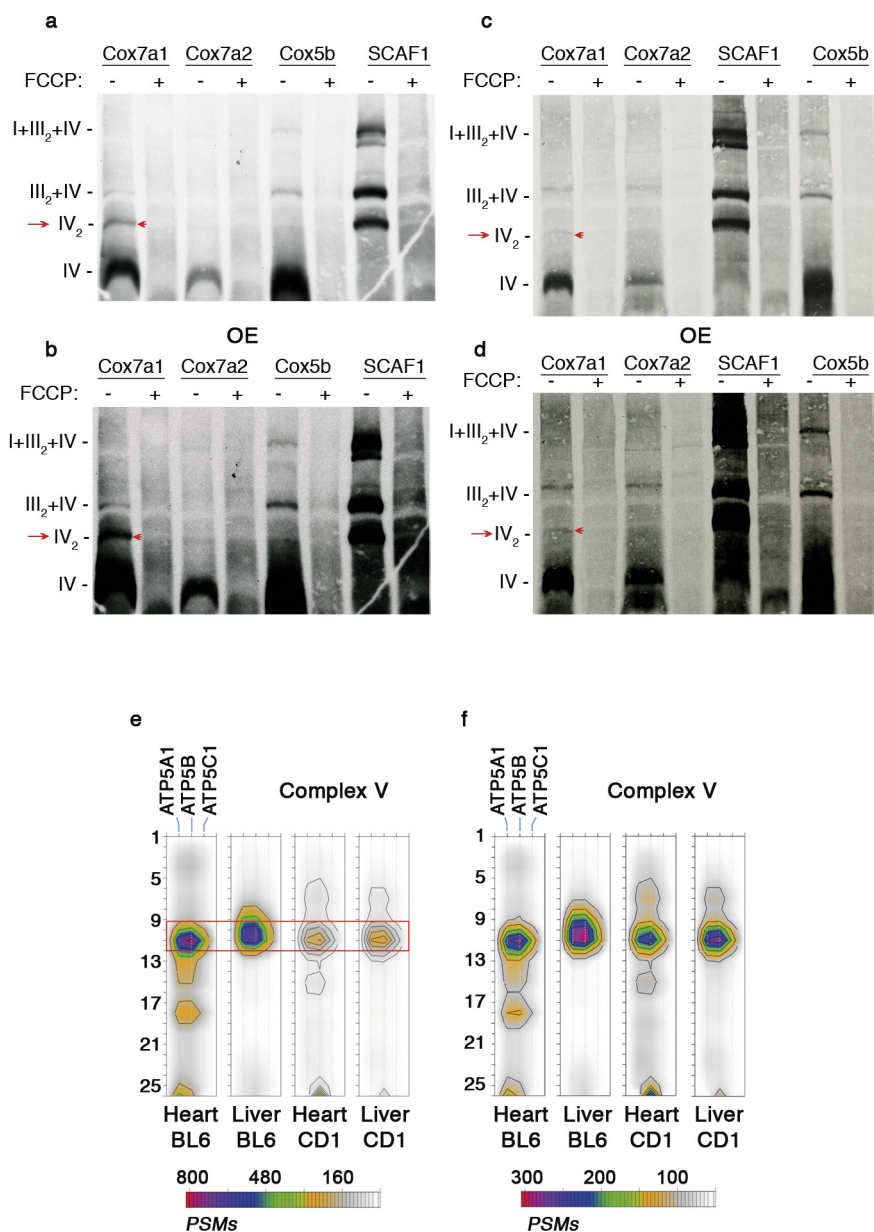
b



Extended Data Figure 3 | Specificity of MS/MS fragmentation series used to generate quantitative peptide profiles for COX6A1 and COX6A2.
a, b, Profiles for COX6A1 (a) and COX6A2 (b). The quality control plots are as in Extended Data Fig. 1.

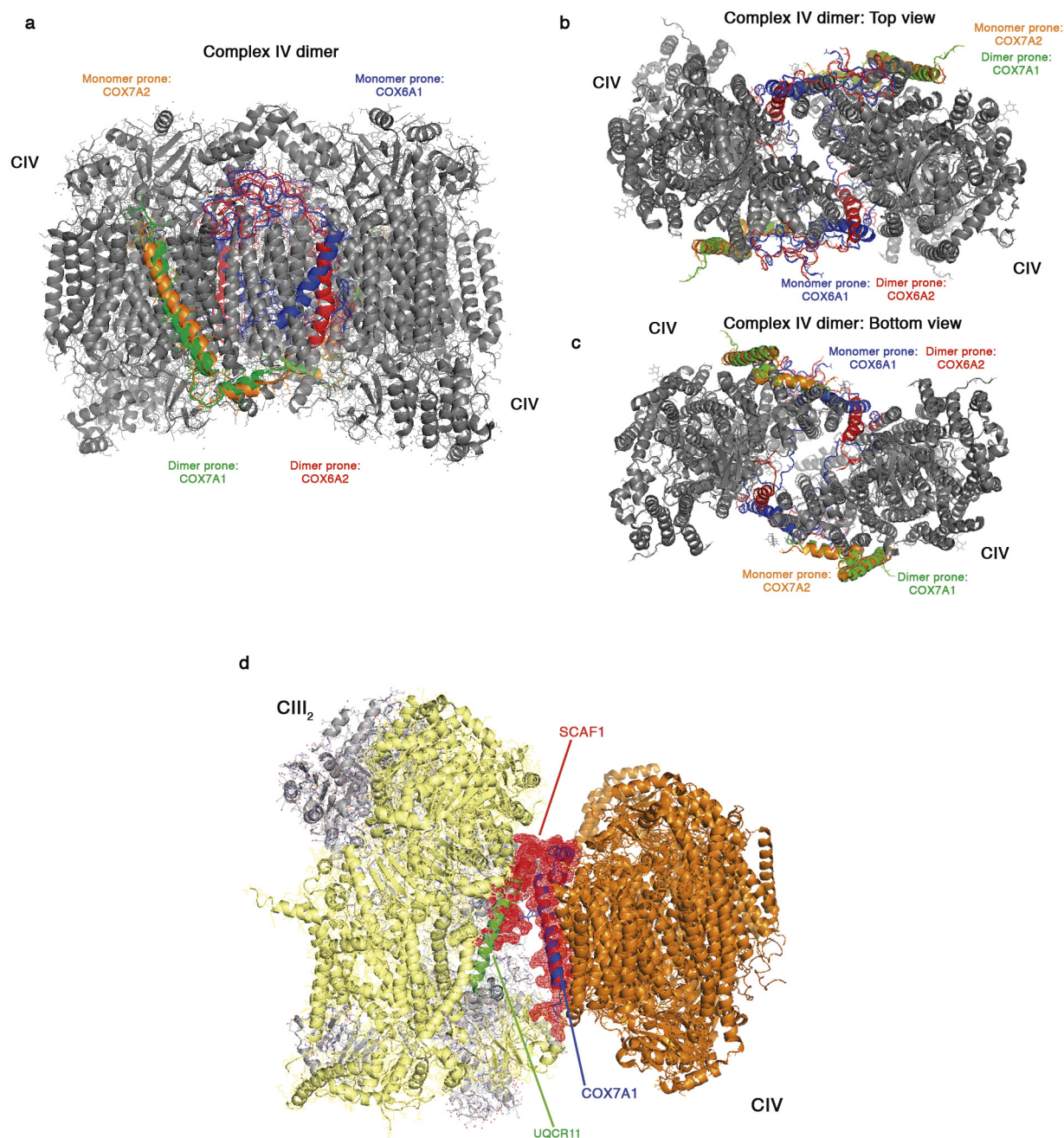


Extended Data Figure 4 | Specificity of MS/MS fragmentation series used to generate quantitative peptide profiles for COX5B and UQCRC1.
a, b, Profiles for COX5B (a) and UQCRC1 (b). The quality control plots are as in Extended Data Fig. 1.



Extended Data Figure 5 | Mitochondrial import assay of CIV proteins and quantitation of CIV by mass spectrometry. **a–d**, Autoradiography of BNGE of CD1 mice liver digitonin-solubilized mitochondria after import of the indicated radiolabelled proteins. Labelling was fully ablated when mitochondrial import was prevented by FCCP-mediated dissipation of the membrane potential. Images represent two biological replicates. OE, overexposed. Red arrows indicate dimer CIV. **e, f**, Quantitative contour

plots of the most abundant proteins of CIV. Data-independent (DiS) mass spectrometry analysis of BNGE gel slices of digitonin-solubilized mitochondria from heart and liver from CD1 and C57BL/6J mice. **e**, Total spectral counts of the indicated components of ATP synthase, illustrating the reproducibility obtained upon slicing of BNGE gels. **f**, Total spectral counts after normalization according to ATP synthase levels.



Extended Data Figure 6 | Structural models of supercomplex assembly.

a–c, Structural modelling of COX6A1 and COX7A2 isoforms on the CIV dimer. The structural model for COX6A1 (in blue) is based on its homology with COX6A2 (PDB 2Y69 chain G, in red) and the structural model for COX7A2 (in orange) is based on its homology with COX7A1 (PDB 2Y69 chain J, in green). **d**, Tentative model showing a possible

supercomplex III₂+IV structure stabilized by a SCAF1 bridge. SCAF1 is proposed to substitute COX7A2 in CIV and to interact with III₂ via its extra domain. The structural model for SCAF1 (red mesh) was generated based on its homology with COX7A1 (PDB 1V54, in blue) and with UQCR11 (PDB 1BGY, in green). CIV is shown in orange and the two components of CIII dimer in yellow and grey.

Cascading MutS and MutL sliding clamps control DNA diffusion to activate mismatch repair

Jiaquan Liu^{1*}, Jeungphil Hanne^{1*}, Brooke M. Britton¹, Jared Bennett¹, Daehyung Kim², Jong-Bong Lee^{2,3} & Richard Fishel^{1,4}

Mismatched nucleotides arise from polymerase misincorporation errors, recombination between heteroallelic parents and chemical or physical DNA damage¹. Highly conserved MutS (MSH) and MutL (MLH/PMS) homologues initiate mismatch repair and, in higher eukaryotes, act as DNA damage sensors that can trigger apoptosis². Defects in human mismatch repair genes cause Lynch syndrome or hereditary non-polyposis colorectal cancer and 10–40% of related sporadic tumours³. However, the collaborative mechanics of MSH and MLH/PMS proteins have not been resolved in any organism. We visualized *Escherichia coli* (Ec) ensemble mismatch repair and confirmed that EcMutS mismatch recognition results in the formation of stable ATP-bound sliding clamps that randomly diffuse along the DNA with intermittent backbone contact. The EcMutS sliding clamps act as a platform to recruit EcMutL onto the mismatched DNA, forming an EcMutS–EcMutL search complex that then closely follows the DNA backbone. ATP binding by EcMutL establishes a second long-lived DNA clamp that oscillates between the principal EcMutS–EcMutL search complex and unrestricted EcMutS and EcMutL sliding clamps. The EcMutH endonuclease that targets mismatch repair excision only binds clamped EcMutL, increasing its DNA association kinetics by more than 1,000-fold. The assembly of an EcMutS–EcMutL–EcMutH search complex illustrates how sequential stable sliding clamps can modulate one-dimensional diffusion mechanics along the DNA to direct mismatch repair.

Mismatch repair (MMR) is an excision-resynthesis system that is initiated at a DNA strand break, which may be located 3' or 5' and hundreds to thousands of base pairs away from the mismatch¹. The fidelity of replication-associated MMR depends on communicating mismatch recognition to the distant strand break and then exclusively directing excision of the strand containing the misincorporation error¹. Resynthesis of the single-strand DNA (ssDNA) gap appears to be independent of excision and is performed by the replicative polymerase¹. Although the source of the strand break that directs MMR in most organisms remains unclear^{1,4–6}, a subset of γ -proteobacteria that includes *E. coli* have evolved DNA adenine methylation (Dam) and the MutH endonuclease to specifically introduce a scission into a transiently hemimethylated Dam GATC site on the newly replicated strand⁷.

How MSH and MLH/PMS proteins communicate mismatch recognition to distant MMR components such as MutH continues to be uncertain⁸. A recent study with *Thermus aquaticus* (Ta) proteins suggested that TaMutL traps TaMutS at the mismatch, which then hypothetically acts as a catalyst to polymerize multiple TaMutL along the DNA⁹. Other studies have visualized *Saccharomyces cerevisiae* (Sc) and human (Hs) MSH and MLH/PMS proteins diffusing along the DNA, although the role of these operations in MMR is unknown^{10–13}. These mechanisms contrast a traditional model where the assembly of a MutS–MutL complex executes directional ATP-hydrolysis-driven DNA translocation to activate distant MMR^{14,15}.

We used single-molecule total internal reflection fluorescence microscopy to image *E. coli* MMR on fields of ~200 well-defined doubly-tethered 17.3-kb mismatched DNAs¹⁶ (Extended Data Fig. 1, Extended Data Table 1, Methods). Injection of Alexa647

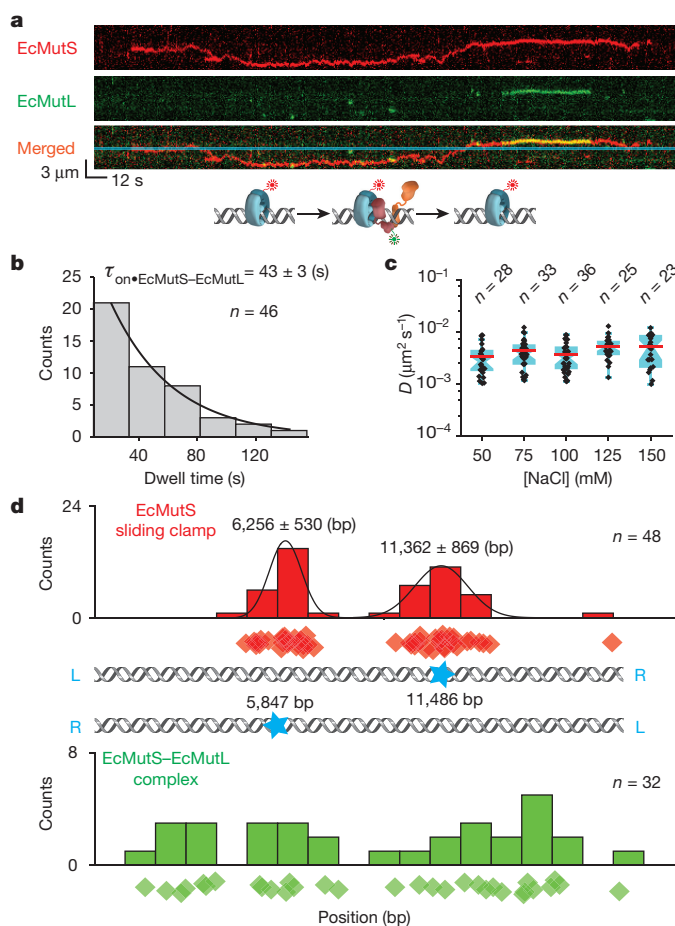


Figure 1 | The formation of an EcMutS–EcMutL complex alters the diffusion properties of EcMutS. **a**, Representative kymographs and illustration showing an EcMutL loaded by an EcMutS sliding clamp. Blue line indicates the mismatch position. **b**, The distribution of dwell times for the EcMutS–EcMutL complex ($\tau_{\text{on-EcMutS-EcMutL}}$; mean \pm s.e.m.). **c**, Box plots of D for the EcMutS–EcMutL complex at different NaCl concentrations (n = number of molecules; Methods). **d**, The distribution of the starting positions for EcMutS (top) or EcMutS–EcMutL complexes on DNA (bottom). There are two possible orientations of the mismatched DNA with mismatch position (blue star; middle panel). Diamonds in top and bottom panels represent individual starting events. Gaussian fit to the top panel distributions is shown as black lines with the mean \pm s.d.

¹Department of Cancer Biology and Genetics, The Ohio State University Wexner Medical Center, Columbus, Ohio 43210, USA. ²Department of Physics, Pohang University of Science and Technology (POSTECH), Pohang, Kyungbuk 790-784, Korea. ³School of Interdisciplinary Bioscience and Bioengineering, POSTECH, Pohang, Kyungbuk, 790-784, Korea. ⁴Department of Physics, The Ohio State University, Columbus, Ohio 43210, USA.

*These authors contributed equally to this work.

(AF647)-labelled EcMutS with ATP¹⁶ (Extended Data Tables 1, 2, Extended Data Fig. 2a, Methods) resulted in numerous particles (93% of total) that diffused along the DNA with a diffusion coefficient ($D_{\text{EcMutS}} = 0.043 \pm 0.016 \mu\text{m}^2 \text{s}^{-1}$) and an association lifetime ($\tau_{\text{on} \cdot \text{EcMutS}} = 185 \pm 35 \text{ s}$) that was two- to three-fold less than that of TaMutS or HsMSH2–HsMSH6 (Extended Data Tables 3, 4, Extended Data Fig. 2c–e)^{13,16,17}. These observations support previous studies showing that mismatch binding provokes the formation of ATP-bound MSH sliding clamps that freely diffuse on DNA^{10,11,13,18,19}. In contrast to other single-molecule imaging studies^{11,20}, we found that Cy3-labelled EcMutL¹⁶ (Extended Data Tables 1, 2, Extended Data Fig. 2a) only binds DNA at low ionic strength or in the absence of magnesium, which was not altered with ATP (Extended Data Fig. 3).

Co-injection resulted in frequent co-localization of EcMutS–AF647 and EcMutL–Cy3 particles on the mismatched DNA (Fig. 1a, Extended Data Fig. 4a, Supplementary Video 1). Most particles (80%) displayed unambiguous real-time coordinated movement along the DNA (Extended Data Table 3, Extended Data Fig. 4b) that permitted determination of the lifetime ($\tau_{\text{on} \cdot \text{EcMutS–EcMutL}} = 43 \pm 3 \text{ s}$) and diffusion coefficient ($D_{\text{EcMutS–EcMutL}} = 0.004 \pm 0.002 \mu\text{m}^2 \text{s}^{-1}$; Fig. 1b, c, Extended Data Table 4). The diffusion coefficient was constant over a range of ionic strengths (Fig. 1c, Extended Data Table 4), indicating the EcMutS–EcMutL complex remains in continuous DNA contact during diffusion²¹. By estimating the Stokes radius from published structures²² we calculated an average free-energy barrier (ϵ) of 1.7 $k_B T$ for rough-landscape diffusion²³ (see Methods). These results are consistent with the conclusion that the EcMutS–EcMutL complex can engage in an efficient DNA search that includes rotation-coupled one-dimensional (1D) diffusion along the backbone^{10,23}.

Although EcMutS sliding clamps originated at the mismatch (Fig. 1d, top), the initial association with EcMutL appeared random along the DNA (Fig. 1d, bottom). Following the assembly of an EcMutS–EcMutL complex, four types of dissociation mechanics were observed (Fig. 2a).

The least common type (2% of total events) was simultaneous dissociation of both EcMutS and EcMutL from the mismatched DNA (Fig. 2a). The most common type (49%) engaged in a dynamic dissociation–association (oscillating) complex on the DNA (Fig. 2a, b, white arrows, Extended Data Fig. 4c, Supplementary Videos 2, 3). The association time for the oscillating EcMutS–EcMutL complex ($\tau_{\text{on} \cdot \text{EcMutS} \leftrightarrow \text{EcMutL}} = 30 \pm 3 \text{ s}$; Extended Data Fig. 4d) was similar to that of the initial EcMutS–EcMutL complex assembly (Fig. 1b). However, the dissociation lifetime was ~ 8 -fold shorter ($\tau_{\text{off} \cdot \text{EcMutS} \leftrightarrow \text{EcMutL}} = 3.8 \pm 0.3 \text{ s}$; Fig. 2c), indicating that EcMutS and EcMutL were in complex $\sim 90\%$ of the time (Fig. 2d). The location of the oscillating EcMutS–EcMutL complexes occurred over the entire DNA length (Fig. 2b, Extended Data Fig. 4c). Interestingly, the EcMutL clamps appeared to bypass the EcMutS clamps in 12% of the oscillation events ($n = 67$; Fig. 2b, lower white arrows).

A third type of EcMutS–EcMutL complex dissociation (22%) left EcMutS alone on the mismatch DNA that appeared to retain identical properties to the initially loaded EcMutS sliding clamps (Fig. 2a), whereas a fourth type (27%) left solitary extremely stable EcMutL particles on the DNA ($\tau_{\text{on} \cdot \text{EcMutL}} = 851 \pm 155 \text{ s}$; Fig. 2e, f, Extended Data Fig. 4e) that diffused ~ 20 times faster than EcMutS sliding clamps and ~ 200 times faster than the EcMutS–EcMutL complex ($D_{\text{EcMutL}} = 0.888 \pm 0.393 \mu\text{m}^2 \text{s}^{-1}$; Fig. 2g, Extended Data Table 4, Supplementary Video 4). D_{EcMutL} increased with ionic strength, indicating that EcMutL diffuses with intermittent DNA contact²¹ (Fig. 2g, Extended Data Table 4). Both EcMutS and ATP were required to observe EcMutS–EcMutL complexes and stable fast-diffusing EcMutL particles on the mismatched DNA (Fig. 2h, Methods).

The ATP dependence, stability and intermittent DNA contact suggest that the EcMutL particles could engender a second sliding clamp. To test this hypothesis, we examined MMR-defective mutant protein EcMutL(R95F), which is incapable of binding ATP

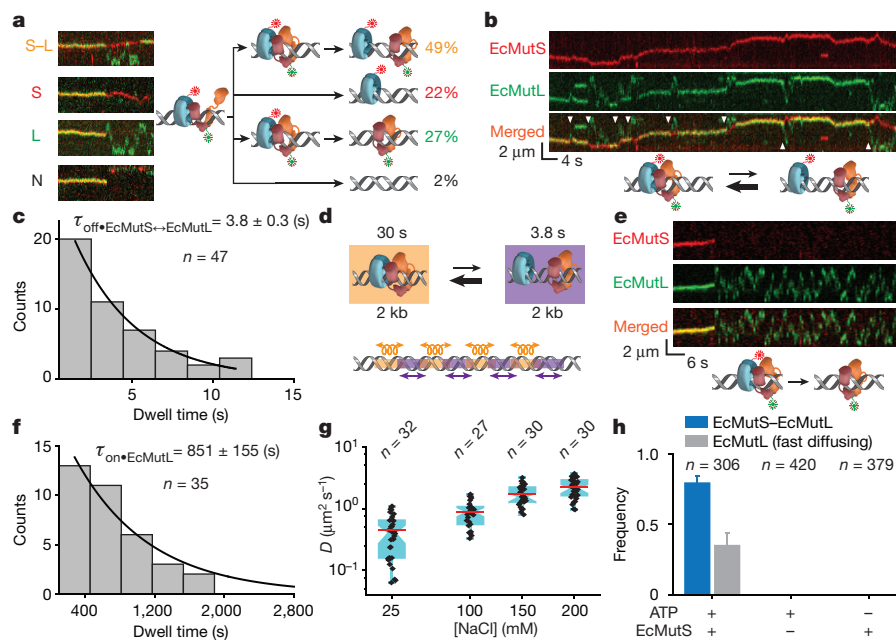


Figure 2 | The formation of an oscillating EcMutS–EcMutL complex and fast-diffusing EcMutL. **a**, Representative kymographs and illustrations showing the different types of EcMutS–EcMutL complex dissociations (coloured letters). The frequency (%) of each dissociation type is shown (right, coloured numbers). **b**, Representative kymographs and illustration showing an oscillating EcMutS–EcMutL complex (white arrowheads indicates dissociation events). **c**, Dissociation time distribution ($\tau_{\text{off} \cdot \text{EcMutS} \leftrightarrow \text{EcMutL}}$; mean \pm s.e.m.) for the oscillating EcMutS–EcMutL complex. **d**, Illustration of oscillating EcMutS–EcMutL

complex with lifetimes and calculated diffusion distances. Oscillations are indicated for illustration only and should be stochastic. **e**, Representative kymographs and illustration showing the dissociation and fast diffusion of EcMutL from an EcMutS–EcMutL complex. **f**, Dwell time distribution (mean \pm s.e.m.) of fast-diffusing EcMutL on the mismatched DNA. **g**, Box plots of D for fast-diffusing EcMutL at different NaCl concentrations. **h**, The frequency of EcMutS–EcMutL complex and fast-diffusing EcMutL under various conditions (mean \pm s.d.); n = number of DNA molecules.

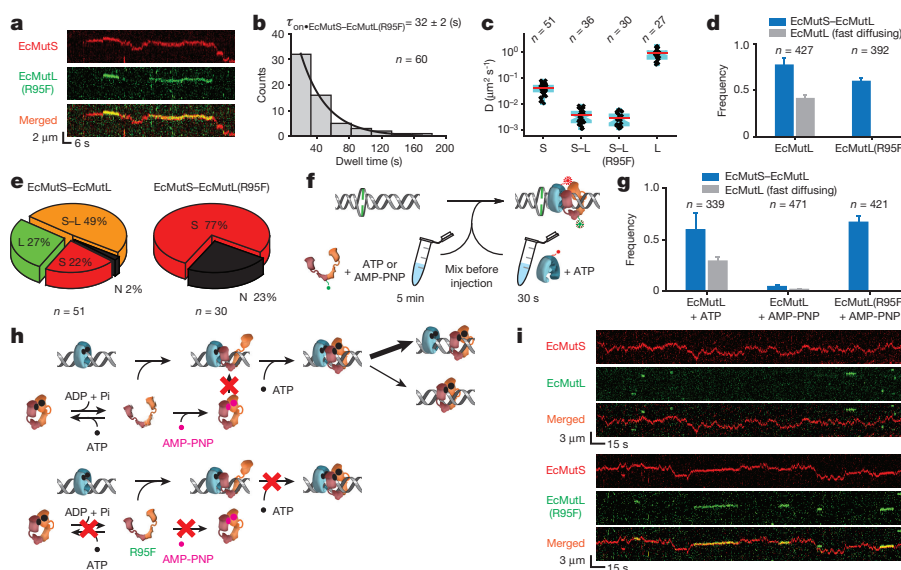


Figure 3 | ATP binding by EcMutL results in formation of a ring-like sliding clamp. **a**, Representative kymographs showing the formation of two separate EcMutS–EcMutL(R95F) complexes on a mismatched DNA. **b**, Dwell time distribution for the EcMutS–EcMutL(R95F) complex ($\tau_{\text{on}}^{\text{EcMutS-EcMutL(R95F)}}$; mean \pm s.e.m.). **c**, Box plots of D for the EcMutS sliding clamp (S); EcMutS–EcMutL complex (S–L); EcMutS–EcMutL(R95F) complex (S–L(R95F)); and EcMutL particle (L) at 100 mM NaCl. **d**, The frequency of EcMutS–EcMutL complex (blue) and fast-diffusing EcMutL (grey) with wild-type EcMutL or EcMutL(R95F) (mean \pm s.d.). **e**, Pie charts showing the distributions of dissociation types

for EcMutS–EcMutL or EcMutS–EcMutL(R95F) complexes. Frequency and dissociation types are indicated within the pie chart (see Fig. 2a). **f**, Illustration of ATP and AMP-PNP pre-incubation studies. **g**, The frequency of EcMutS–EcMutL complex and fast-diffusing EcMutL under various EcMutL or EcMutL(R95F) pre-incubation conditions (mean \pm s.d.). **h**, Illustration showing AMP-PNP induced EcMutL ring-like clamp closure (top) that does not affect EcMutL(R95F) (bottom). **i**, Representative kymographs showing the absence of EcMutS–EcMutL complexes and the presence of EcMutS–EcMutL(R95F) complexes following pre-incubation of EcMutL or EcMutL(R95F) with AMP-PNP.

or activating EcMutH^{18,24} (Extended Data Tables 1, 2, Extended Data Fig. 2a, b). Numerous EcMutS–EcMutL(R95F) complexes were observed on the mismatched DNA that displayed a lifetime ($\tau_{\text{on}}^{\text{EcMutS-EcMutL(R95F)}} = 32 \pm 2$ s) and diffusion ($D_{\text{EcMutS-EcMutL(R95F)}} = 0.003 \pm 0.002 \mu\text{m}^2 \text{s}^{-1}$) similar to the wild-type EcMutS–EcMutL complex (85% of the particles; Fig. 3a–c, Extended Data Tables 3, 4). These results confirm that EcMutL(R95F) associates normally with EcMutS sliding clamps¹⁸. However, oscillating EcMutS–EcMutL(R95F) and stable fast-diffusing EcMutL(R95F) were never observed (Fig. 3d, Methods), with the majority of complex dissociations leaving EcMutS alone on the DNA (77%, Fig. 3e). We conclude that a principal role for ATP binding by EcMutL is to establish an oscillating EcMutS–EcMutL complex and/or stable fast-diffusing EcMutL.

Pre-binding EcMutS with a poorly hydrolysable ATP analogue inhibits sliding clamp formation¹⁸, because it induces a closed ring-like structure that without hydrolysis is unable to interact with the mismatched DNA^{22,25}. Similarly, a non-hydrolysable ATP analogue bound to EcMutL might produce a closed ring-like structure that is incapable of interacting with EcMutS or forming stable EcMutL particles on the mismatched DNA (Fig. 3f). As a control, pre-incubation of EcMutL with ATP did not significantly change the frequency of EcMutS–EcMutL complexes or stable EcMutL particles on the mismatched DNA (Fig. 3g, left; compared to Fig. 2h, Fig. 3d, Methods). However, pre-incubation of EcMutL with non-hydrolysable adenylyl-imidodiphosphate (AMP-PNP) substantially reduced the frequency of EcMutS–EcMutL complexes and stable EcMutL particles (Fig. 3g, middle, Fig. 3h, top), leaving mostly EcMutS sliding clamps on the DNA (Fig. 3i, top). Because EcMutL(R95F) cannot bind ATP, abundant EcMutS–EcMutL(R95F) complexes were observed following pre-incubation of AMP-PNP, which then never progressed to oscillating EcMutS–EcMutL or stable EcMutL particles (Fig. 3g, right, Fig. 3h, bottom, and Fig. 3i, bottom). These results are consistent with the idea that an open conformation of EcMutL is required to interact with EcMutS sliding clamps, which then binds ATP to form a second, exceedingly stable, ring-like clamp.

The EcMutH dissociation constant (K_d) with hemimethylated Dam GATC DNA is at least threefold higher than its cellular concentration^{26,27}. EcMutL activation of the EcMutH endonuclease might overcome this binding problem^{18,28}, although the mechanics of this progression are unknown. We only detected AF647-labelled EcMutH (Extended Data Table 1, Extended Data Figs 2a, 5a) bound to the mismatched DNA in ≤ 50 mM NaCl (Extended Data Fig. 5b–g). However, when both EcMutS and EcMutL were present, numerous EcMutH particles were observed on the mismatched DNA at physiological ionic concentration (Fig. 4a, Methods). The majority of EcMutH–AF647 (89% of particles) co-localized and displayed coordinated diffusion with EcMutL–Cy3 (Fig. 4b, Extended Data Table 3). No EcMutH binding was observed when EcMutL(R95F) replaced wild-type EcMutL (Fig. 4a, c), suggesting that EcMutH only associates with an ATP-bound EcMutL sliding clamp.

By substituting EcMutS–AF555 for EcMutL–Cy3, we observed EcMutS–EcMutL–EcMutH complexes (Fig. 4d) that oscillated with free EcMutS and EcMutL–EcMutH, and displayed similar ionic-strength-independent diffusion properties to the EcMutS–EcMutL complex ($D_{\text{EcMutS-EcMutL-EcMutH}} = 0.005 \pm 0.003 \mu\text{m}^2 \text{s}^{-1}$; Extended Data Table 4, Extended Data Fig. 5h, middle, Extended Data Fig. 5i, Supplementary Video 5). As predicted, labelled EcMutL–EcMutH complexes (with unlabelled EcMutS) oscillated with a mean diffusion coefficient that was skewed faster than the EcMutS–EcMutL complex ($P = 0.001$; $D_{\text{(EcMutS)-EcMutL-EcMutH}} = 0.010 \pm 0.010 \mu\text{m}^2 \text{s}^{-1}$; Fig. 4e, Extended Data Fig. 5h, left, Extended Data Table 4). The lifetime of the EcMutL–EcMutH complex appeared to be independent of EcMutS association and increased the intrinsic EcMutH DNA binding kinetics by at least 1,000-fold ($\tau_{\text{on}}^{\text{EcMutL-EcMutH}} = 187 \pm 6$ s; compare Fig. 4f to Extended Data Fig. 5c). The diffusion coefficient of free EcMutL–EcMutH complex was intermediate between EcMutS–EcMutL complex and EcMutL sliding clamp ($P < 0.0001$; $D_{\text{EcMutL-EcMutH}} = 0.041 \pm 0.034 \mu\text{m}^2 \text{s}^{-1}$; Extended Data Fig. 5h, right; Methods) and increased with ionic strength, indicating that the EcMutL–EcMutH complex moves with intermittent DNA contact²¹

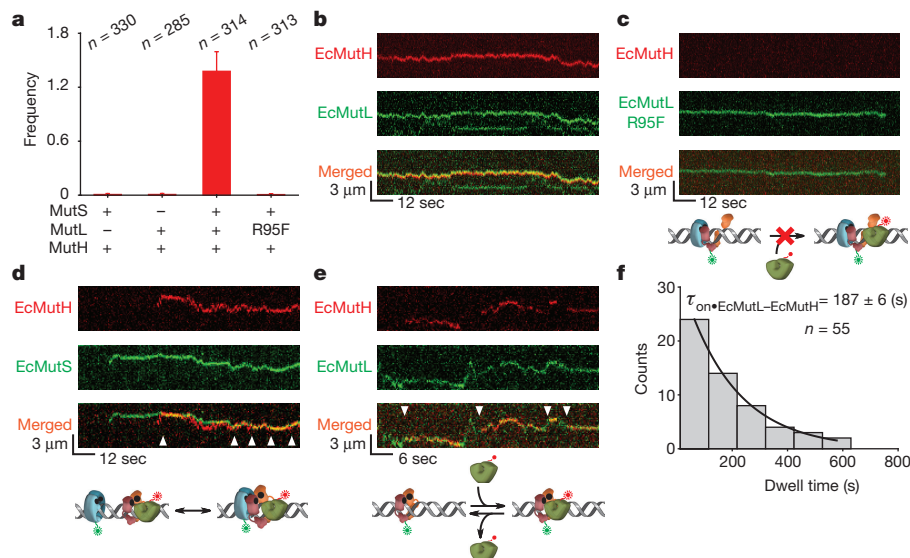


Figure 4 | EcMutH binds EcMutL sliding clamps. **a**, The frequency of EcMutH-AF647 on the mismatched DNA in the presence of other MMR components (mean \pm s.d.). **b**, Representative kymographs of EcMutH-AF647 co-localization and diffusion with EcMutL-Cy3 (and unlabelled EcMutS) on a single mismatched DNA. **c**, Representative kymographs of EcMutL(R95F)-Cy3 (and unlabelled EcMutS) diffusion on a single mismatched DNA. Note the absence of EcMutH-AF647. **d**, Representative kymographs of EcMutH-AF647 co-localization and diffusion with

EcMutS-AF555 (and unlabelled EcMutL) on a single mismatched DNA. Arrowheads indicate association of EcMutS with EcMutL-EcMutH. **e**, Representative kymographs of EcMutH-AF647 co-localization and diffusion with EcMutL-Cy3 on a single mismatched DNA following the dissociation of EcMutS. Arrowheads indicate association of EcMutH with EcMutL sliding clamp. **f**, Dwell time distribution of the EcMutL-EcMutH complex ($\tau_{\text{on}} \cdot \text{EcMutL-EcMutH}$; mean \pm s.e.m.).

(Extended Data Fig. 5j, Extended Data Table 4). Together these observations suggest that the EcMutS-EcMutL-EcMutH complex is uniquely capable of engaging in an efficient 1D diffusion-mediated DNA search for a hemimethylated Dam GATC site.

Our studies demonstrate that *E. coli* MMR employs a cascade of stable ATP-bound sliding clamps to modulate 1D diffusion mechanics along the DNA. These observations show how mismatch recognition is effectively communicated to a distant MMR initiation site and provides additional support for the molecular switch/sliding clamp model^{18,29}. In this model, EcMutS mismatch recognition forms ATP-bound sliding clamps that randomly diffuse with intermittent DNA contact, establishing a stable platform capable of attracting other MMR components^{10,13,18,19,30} (Extended Data Fig. 6a). The association of EcMutL with an EcMutS sliding clamp creates a previously unrecognized EcMutL sliding clamp that displays binding oscillation with EcMutS (Extended Data Fig. 6b). Both an EcMutS sliding clamp and ATP binding by EcMutL are required to create these configurations (Extended Data Fig. 6b). Single-molecule images of quantum-dot-labelled ScMsh2-ScMsh6 and ScMlh1-ScPms1 on duplex²⁰ and mismatched DNA curtains¹¹ appeared unable to detect these conformations (Supplementary Note 1).

We confirmed that the N terminus of EcMutL interacts with EcMutS sliding clamps using Förster resonance energy transfer²² (FRET; Extended Data Fig. 5k, Supplementary Video 6). These observations are consistent with a molecular progression where one of the EcMutL N-terminal homodimer domains interacts with EcMutS²², creating a scaffold for the remaining EcMutL domains to wrap around the DNA in concert with natural helical motions during complex diffusion along the backbone; ultimately triggering ATP-dependent N-terminal domain dimerization and clamp formation²⁴. The fast diffusion and intermittent DNA contact of EcMutL sliding clamps makes ATP-induced collapse of the linker arms between the N- and C-terminal domains unlikely³¹. Indeed, the EcMutS sliding clamps were capable of passing through the EcMutL sliding clamps in 12% of binding-oscillation encounters. We conclude that the fundamental role of ATP binding by EcMutL is to form an exceedingly stable sliding clamp on mismatched DNA. This observation might explain the persistence of ScMlh1-ScPms1 during MMR *in vivo*³².

The EcMutS-EcMutL complex remains in continuous DNA contact and capable of performing an efficient 1D diffusion-mediated DNA search²³. The oscillation of EcMutS-EcMutL with free EcMutS and EcMutL sliding clamps suggests that the complex may search ~ 2 kb of DNA, dissociate, move ~ 2 kb, then re-associate and begin a new ~ 2 kb search. This segmented search cycle may interrogate ~ 12 kb of DNA during the limiting lifetime of EcMutS sliding clamps (Extended Data Fig. 6a). Visualizing EcMutS-EcMutL diffusion clearly contrasts a trapped TaMutS-TaMutL⁹. It is likely that the TaMutS-TaMutL studies were influenced by an obligatory FRET constraint near the mismatch and/or the use of thermophilic proteins that become active at higher temperatures.

EcMutH exclusively associated with the ATP-bound EcMutL either in the EcMutS-EcMutL search complex or as a free sliding clamp (Extended Data Fig. 6c). Importantly, the EcMutL-EcMutH complex diffuses with intermittent DNA contact, albeit ~ 20 -fold slower than EcMutL. A simple explanation is that EcMutH increases the sporadic DNA interactions slowing EcMutL-EcMutH diffusion along the DNA. Nevertheless, the intermittent DNA contact implies that the EcMutL-EcMutH complex is probably not capable of performing an efficient DNA search. In contrast, the EcMutS-EcMutL-EcMutH complex remains in continuous contact with the DNA backbone allowing it to engage in a 1D diffusion-mediated search that may efficiently locate a distant hemimethylated Dam GATC site (Extended Data Fig. 6d).

A mechanism that utilizes modulated stochastic diffusion along the DNA to initiate excision contrasts most historical MMR models^{8,9,14,15}. The notable conservation of MSH and MLH/PMS proteins implies that a similar mechanics is probably widely used for exploring the DNA backbone and activating distant MMR processes; it might also be used by MSH and MLH/PMS to connect damage sensing to the DNA damage response machinery².

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 19 May; accepted 21 October 2016.

Published online 16 November 2016.

1. Fishel, R. Mismatch repair. *J. Biol. Chem.* **290**, 26395–26403 (2015).
2. Wang, J. Y. & Edelmann, W. Mismatch repair proteins as sensors of alkylation DNA damage. *Cancer Cell* **9**, 417–418 (2006).
3. Martín-López, J. V. & Fishel, R. The mechanism of mismatch repair and the functional analysis of mismatch repair defects in Lynch syndrome. *Fam. Cancer* **12**, 159–168 (2013).
4. Ghodgaonkar, M. M. *et al.* Ribonucleotides misincorporated into DNA act as strand-discrimination signals in eukaryotic mismatch repair. *Mol. Cell* **50**, 323–332 (2013).
5. Kawasoe, Y., Tsurimoto, T., Nakagawa, T., Masukata, H. & Takahashi, T. S. MutS α maintains the mismatch repair capability by inhibiting PCNA unloading. *eLife* **5**, e15155 (2016).
6. Lujan, S. A., Williams, J. S., Clausen, A. R., Clark, A. B. & Kunkel, T. A. Ribonucleotides are signals for mismatch repair of leading-strand replication errors. *Mol. Cell* **50**, 437–443 (2013).
7. Putnam, C. D. Evolution of the methyl directed mismatch repair system in *Escherichia coli*. *DNA Repair (Amst.)* **38**, 32–41 (2016).
8. Kolodner, R. D., Mendillo, M. L. & Putnam, C. D. Coupling distant sites in DNA during DNA mismatch repair. *Proc. Natl Acad. Sci. USA* **104**, 12953–12954 (2007).
9. Qiu, R. *et al.* MutL traps MutS at a DNA mismatch. *Proc. Natl Acad. Sci. USA* **112**, 10914–10919 (2015).
10. Cho, W. K. *et al.* ATP alters the diffusion mechanics of MutS on mismatched DNA. *Structure* **20**, 1264–1274 (2012).
11. Gorman, J. *et al.* Single-molecule imaging reveals target-search mechanisms during DNA mismatch repair. *Proc. Natl Acad. Sci. USA* **109**, E3074–E3083 (2012).
12. Jeon, Y. *et al.* Dynamic control of strand excision during human DNA mismatch repair. *Proc. Natl Acad. Sci. USA* **113**, 3281–3286 (2016).
13. Jeong, C. *et al.* MutS switches between two fundamentally distinct clamps during mismatch repair. *Nat. Struct. Mol. Biol.* **18**, 379–385 (2011).
14. Kunkel, T. A. Celebrating DNA's repair crew. *Cell* **163**, 1301–1303 (2015).
15. Modrich, P. Mechanisms and biological effects of mismatch repair. *Annu. Rev. Genet.* **25**, 229–253 (1991).
16. Liu, J. *et al.* An efficient site-specific method for irreversible covalent labeling of proteins with a fluorophore. *Sci. Rep.* **5**, 16883 (2015).
17. Heinen, C. D. *et al.* Human MSH2 (hMSH2) protein controls ATP processing by hMSH2–hMSH6. *J. Biol. Chem.* **286**, 40287–40295 (2011).
18. Acharya, S., Foster, P. L., Brooks, P. & Fishel, R. The coordinated functions of the *E. coli* MutS and MutL proteins in mismatch repair. *Mol. Cell* **12**, 233–246 (2003).
19. Gradia, S. *et al.* hMSH2–hMSH6 forms a hydrolysis-independent sliding clamp on mismatched DNA. *Mol. Cell* **3**, 255–261 (1999).
20. Gorman, J., Plys, A. J., Visnapuu, M. L., Alani, E. & Greene, E. C. Visualizing one-dimensional diffusion of eukaryotic DNA repair factors along a chromatin lattice. *Nat. Struct. Mol. Biol.* **17**, 932–938 (2010).
21. Berg, O. G. & von Hippel, P. H. Selection of DNA binding sites by regulatory proteins. Statistical-mechanical theory and application to operators and promoters. *J. Mol. Biol.* **193**, 723–750 (1987).
22. Groothuizen, F. S. *et al.* MutS/MutL crystal structure reveals that the MutS sliding clamp loads MutL onto DNA. *eLife* **4**, e06744 (2015).
23. Blainey, P. C. *et al.* Nonspecifically bound proteins spin while diffusing along DNA. *Nat. Struct. Mol. Biol.* **16**, 1224–1229 (2009).
24. Ban, C., Junop, M. & Yang, W. Transformation of MutL by ATP binding and hydrolysis: a switch in DNA mismatch repair. *Cell* **97**, 85–97 (1999).
25. Lamers, M. H., Winterwerp, H. H. & Sixma, T. K. The alternating ATPase domains of MutS control DNA mismatch repair. *EMBO J.* **22**, 746–756 (2003).
26. Lee, J. Y. *et al.* Muth complexed with hemi- and unmethylated DNAs: coupling base recognition and DNA cleavage. *Mol. Cell* **20**, 155–166 (2005).
27. Feng, G., Tsui, H. C. & Winkler, M. E. Depletion of the cellular amounts of the MutS and Muth methyl-directed mismatch repair proteins in stationary-phase *Escherichia coli* K-12 cells. *J. Bacteriol.* **178**, 2388–2396 (1996).
28. Hall, M. C. & Matson, S. W. The *Escherichia coli* MutL protein physically interacts with Muth and stimulates the Muth-associated endonuclease activity. *J. Biol. Chem.* **274**, 1306–1312 (1999).
29. Fishel, R. Mismatch repair, molecular switches, and signal transduction. *Genes Dev.* **12**, 2096–2101 (1998).
30. Gradia, S., Acharya, S. & Fishel, R. The human mismatch recognition complex hMSH2–hMSH6 functions as a novel molecular switch. *Cell* **91**, 995–1005 (1997).
31. Sacho, E. J., Kadyrov, F. A., Modrich, P., Kunkel, T. A. & Erie, D. A. Direct visualization of asymmetric adenine-nucleotide-induced conformational changes in MutL α . *Mol. Cell* **29**, 112–121 (2008).
32. Hombauer, H., Campbell, C. S., Smith, C. E., Desai, A. & Kolodner, R. D. Visualization of eukaryotic DNA mismatch repair reveals distinct recognition and repair intermediates. *Cell* **147**, 1040–1053 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements We would like to thank C. Bell and our laboratory colleagues for many helpful insights and discussions. This work was supported by NRF of Korea Grant No. 2011-001309 (J.-B.L.) and NIH grant CA67007 (R.F.).

Author Contributions J.L., J.H., J.-B.L. and R.F. designed the experiments; B.M.B., and J.B., performed genetic analysis; J.L. purified and labelled the proteins; J.L. and J.H. performed the single-molecule studies; J.L., J.H., J.-B.L. and R.F. analysed the data; D.K. developed the Matlab script for diffusion analysis and particle co-localization. J.L., J.H., J.-B.L. and R.F. wrote the paper and all authors participated in critical discussions.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.-B.L. (jblee@postech.ac.kr) or R.F. (fishel.7@osu.edu).

METHODS

Buffers and experiment conditions. The single-molecule imaging Buffer A contains 20 mM Tris-HCl (pH 7.5), 5 mM MgCl₂, 0.1 mM DTT, 0.2 mg ml⁻¹ acetylated BSA (Promega), 0.0025% P-20 surfactant (GE healthcare) and 100 mM NaCl (unless stated otherwise). To minimize photoblinking and photobleaching, imaging buffer was supplemented with a photostability enhancing and oxygen scavenging cocktail containing saturated (~3 mM) Trolox and PCA/PCD oxygen scavenger system composed of PCA (1 mM) and PCD (0.42 U ml⁻¹, rPCO, OY Americas)³³.

Construction of 17.3-kb λ-phage-based DNA with a single mismatch. λ-phage DNA (3.2 nM, Thermo Scientific) was ligated with the lambda mismatch 1 oligonucleotide (800 nM; Extended Data Fig. 1, Extended Data Table 1) at room temperature (22°C) overnight. The unligated oligonucleotides were removed by a 100 kDa Amicon filter (Millipore). The resulting DNA (3.2 nM) was then ligated with lambda mismatch 2 (800 nM) at room temperature overnight, followed by the removal of unligated oligonucleotides as above. The resulting DNA (1.6 nM) was cyclized at 18°C overnight and digested with BsaI at 37°C for 4 h. The BsaI-treated DNA was then ligated with the lambda linkers (~2 μM, Extended Data Table 1) at 18°C overnight. The DNA ligation products were separated on a 0.5% low melting agarose (Promega) gel; the 17.3-kb band was excised and treated with β-agarase (New England Biolabs) followed by isopropanol precipitation. The purified DNA was resuspended in TE buffer (10 mM Tris-HCl, pH 7.5, 1 mM EDTA) and stored at -20°C until use. To methylate the mismatched DNA (Fig. 4, Extended Data Fig. 5h–j), 1 μg mismatched DNA was incubated with 80 μM S-adenosylmethionine and 8 U of Dam methyltransferase (New England Biolabs) at 37°C for 2 h in an 100-μl reaction, followed by inactivation of the enzyme at 65°C for 15 min. Control experiments demonstrate that the DNA becomes completely resistant to MboI, indicating full methylation.

Plasmid construction, MMR protein labelling and purification. EcMutS and EcMutL expression constructs have been previously described¹⁶. The EcMutL(R95F) point mutation was generated using the QuikChange site-directed mutagenesis kit (Stratagene). Hexa-histidine (his₆) and formylglycine-generating enzyme (FGE) recognition hexa-amino acid sequence (LCTPSR; ald₆) were introduced on to the C terminus of all MMR proteins with the exception of EcMutL(N-ter) and EcMutH, in which the tags were introduced on to the N terminus using a previously described protocol¹⁶. Two glycine residues separated the his₆ and ald₆ from one another and these tags were separated from the MMR proteins by two serine residues. The order of these tags relative to the MMR gene is indicated in Extended Data Table 2. The *E. coli mutH* gene was amplified by PCR (Extended Data Table 1), digested with NdeI and BamHI, and inserted into pET-29a (Novagen) bacterial expression plasmid. The resulting construct was amplified in *E. coli* XL10 gold (Stratagene) and verified by DNA sequencing. All the EcMutS and EcMutL proteins were expressed, labelled and purified as described previously¹⁶. EcMutH was expressed, labelled and purified by a protocol similar to EcMutS. All of the his₆- and ald₆-tagged MMR gene constructs used in these studies were shown to have a wild-type phenotype by complementation analysis in respective MMR-deficient *E. coli* strains except EcMutL(R95F), which was genetically a mutator indicating an MMR-deficiency as a result of a complete inability to bind ATP¹⁸ (Extended Data Table 2). We used EcMutS(D835R,R840E) since the combined D835R and R840E substitution mutations eliminate the anomalous confounding effects of EcMutS protein tetramerization *in vitro*³⁴ (Extended Data Table 2). The labelling efficiencies of EcMutS-AF647 dimer (45%), EcMutL-Cy3 dimer (33%), EcMutL(R95F)-Cy3 dimer (29%) and EcMutH-AF647 monomer (50%) were determined by spectrophotometry as described¹⁶ (Extended Data Table 1).

Total internal reflection fluorescence microscopy and single-molecule studies. All the single-molecule data in this study were acquired on a custom-built prism-type total internal reflection fluorescence microscope based on the Olympus microscope body IX71. Fluorophores were excited using the laser lines (532 nm for green, 635 nm for red) in the single-molecule total internal reflection fluorescence microscopy system. Image acquisition was performed using an EMCCD camera (ProEM Exelon512, Princeton Instruments) after splitting emissions by a Dual View optical setup (DV2, Photometrics). Micro-Manager image capture software was used to control the opening and closing of a shutter, which in turn controlled the laser excitation³⁵.

The 17.3-kb mismatched DNA (300 pM) in 300 μl T50 buffer (20 mM Tris-HCl, pH 7.5, 50 mM NaCl) was injected into the flow cell chamber and stretched by laminar flow (250 μl min⁻¹). The stretched DNA was anchored onto a streptavidin-coated, PEG passivated quartz slide surface, and the unbound DNA was flushed by similar laminar flow.

To detect transiently bound EcMutL or EcMutH on DNA (Extended Data Figs 3, 5b–g), EcMutL or EcMutH (2–10 nM) in imaging buffer was introduced into the flow cell chamber and protein-DNA interactions were monitored in real-time in

the absence of flow at ambient temperature. The DNA was stained with Sytox Orange (250 nM, Invitrogen) or Syto 59 (700 nM, Invitrogen) after recording.

To measure the interaction between EcMutS and EcMutL, EcMutS (10 nM) and/or EcMutL (20 nM) in imaging buffer plus 1 mM ATP (unless stated otherwise) were introduced into the flow cell chamber and protein-DNA interactions were monitored.

Fully methylated DNA was used in the EcMutS-EcMutL-EcMutH experiments (Fig. 4). To measure the interaction between EcMutL and EcMutH, unlabelled EcMutS (5 nM), EcMutL-Cy3 (10 nM) and EcMutH-AF647 (5 nM) in imaging buffer plus 1 mM ATP were introduced. To measure the interaction between EcMutS and EcMutH, EcMutS-AF555 (10 nM), unlabelled EcMutL (20 nM) and EcMutH-AF647 (5 nM) in imaging buffer plus 1 mM ATP were introduced.

To determine the EcMutL sliding clamp diffusion coefficient, EcMutS and EcMutL in imaging buffer (100 mM NaCl) plus 1 mM ATP were first co-injected. After 5 min the flow cell was flushed with imaging buffer containing 25–200 mM NaCl and 1 mM ATP to remove unbound proteins and EcMutL sliding clamps were monitored.

To determine the diffusion coefficient of the free EcMutL-EcMutH complex, unlabelled EcMutS (5 nM) and EcMutL-Cy3 (10 nM) in imaging buffer (100 mM NaCl) plus 1 mM ATP were first co-injected. After 5 min the flow cell was flushed with imaging buffer containing 100 mM NaCl and 1 mM ATP to remove unbound proteins. EcMutH-AF647 (5 nM) in imaging buffer containing 50–150 mM NaCl and 1 mM ATP was then introduced. 15 min after the wash diffusing EcMutH particles were monitored.

AMP-PNP pre-incubation. EcMutL or EcMutL(R95F) (200 nM) was incubated with 2.5 mM ATP or AMP-PNP in 100 μl Buffer A for 5 min. In a separate reaction, EcMutS (11 nM) was incubated with 1.1 mM ATP in 900 μl Buffer A for 30 s. The two reactions were then mixed and introduced into the flow cell chamber immediately followed by the single-molecule imaging.

Single-molecule data analysis. We found that 93% (52 out of 56) of the EcMutS sliding clamps displayed 1D diffusion along the mismatch DNA, whereas 7% (4 out of 56) remained on the mismatch (Extended Data Table 3). It is likely that the 7% EcMutS that remained on the mismatch reflects inactive proteins that do not respond to ATP binding^{11,13} and was excluded from further analysis. Molecules that diffused for at least 10 s (300–1,500 ms frame rate) were included in the diffusion analysis.

For studies involving EcMutS and EcMutL, fluorescent molecules in two channels were co-localized using a custom written MATLAB script. In these studies 80% (43 out of 54) of the co-localized molecules displayed 1D diffusion, while 20% (11 out of 54) exhibited insignificant movement (Extended Data Table 3). Most of the stationary molecules appeared to be associated with the flow-cell surface and not subsequently stained DNA molecules (Extended Data Table 3). A similar distribution was observed for the EcMutS-EcMutL(R95F) complex (Extended Data Table 3). We excluded these immobile molecules from further analysis and only included molecules that diffused for at least 10 s (250–600 ms frame rate).

In the presence of EcMutS and EcMutL, 89% (85 out of 95) of the EcMutH molecules displayed stable 1D diffusion along the DNA, while 11% (10 out of 95) exhibited insignificant movement (Extended Data Table 3). Most of the stationary molecules appeared to be associated with the flow-cell surface and not subsequently stained DNA molecules (Extended Data Table 3). We excluded these immobile molecules from further analysis and only included molecules that diffused for at least 10 s.

EcMutL sliding clamps were easily distinguished from the EcMutS-EcMutL complex as a result of its extremely fast diffusion along the DNA (Fig. 2e, g). We included molecules with $D > 0.05 \mu\text{m}^2 \text{s}^{-1}$ for this analysis.

Position determination on DNA. After the real-time measures, the 17.3-kb mismatched DNA was stained with Syto 59 (700 nM, Invitrogen) or Sytox Orange (250 nM, Invitrogen). The left (P_L) and the right (P_R) end positions of the DNA were determined by plotting the fluorescent intensities along the length of the stained DNA (Extended Data Fig. 1d). Horizontal positions of diffusing particles (P_P) along the length of the DNA were tracked by DiaTrack 3.04 (Sydney, Australia), in which the particle intensities were fit to a two-dimensional Gaussian function to obtain their positions with sub-pixel resolution. The positions were then converted to lengths in bp by the following equation: $17,332 \text{ bp} \times (P_P - P_L) / (P_R - P_L)$, where 17,332 bp is the length of the mismatched DNA. A 1,000 bp (~2 pixels) binning size was used to construct the position histograms.

Diffusion coefficient measurement and diffusion distance calculation. Particles were tracked using DiaTrack as described above to obtain single-molecule trajectories. Diffusion coefficients were calculated from the trajectories as described previously³⁶. Briefly, the diffusion coefficient (D) was determined from the slope of a mean-square displacement (MSD) versus time plot using the equation $\text{MSD}(t) = 2Dt$, where t is the time interval. The first 10% of the

total measurement time was taken for point fitting. A minimum number of 50 frames were used to calculate the diffusion coefficients except for EcMutS–EcMutL complexes at 125 mM or 150 mM NaCl, in which 30 frames were used as the minimum because of the shortened lifetime. The MSD of each molecule was calculated by:

$$\text{MSD}(t) = \frac{1}{N} \sum_{i=1}^N (x_i(t) - x_i(0))^2$$

where $t = n\Delta T$ ($n = 1, 2, 3, \dots$, ΔT is the sampling rate), N = total number of diffusion events, x = horizontal position, and i = diffusion event. Diffusion coefficients were graphed in box plots (Figs. 1c, 2g, 3c, Extended Data Fig. 5) showing the mean (red line), median (indentation), the upper and the lower quartiles (box ends) and the outliers (whiskers). Diamonds indicate individual events. The average diffusion distance (d) was calculated using the equation: $d = \sqrt{2D\tau}$, where τ is the average lifetime of the particles on the DNA (Fig. 2d). Only EcMutS diffusion distance was reported during dissociation events of EcMutS–EcMutL complex.

Lifetime measurement. A 100-ms or 200-ms frame rate was used to measure transient-bound EcMutL or EcMutH on DNA. A 1.5-s frame rate with 300-ms laser exposure time was used to minimize photo-bleaching during EcMutS sliding clamp or EcMutL sliding clamp lifetime measurements. A 1.0-s frame rate with 300-ms laser exposure time was used to minimize photo-bleaching during EcMutL–EcMutH complex lifetime measurements. A 600-ms frame rate with a 300-ms laser exposure time was used to detect transient association-dissociation during EcMutS–EcMutL complex lifetime measurements. Kymographs were generated along the DNA by a kymograph plugin in ImageJ (J. Rietdorf and A. Seitz, EMBL Heidelberg). The dwell times of EcMutS sliding clamps, EcMutL particles or EcMutH particles on DNA were obtained from kymographs and binned to generate dwell time distributions, which were then fit to single exponential decay functions to determine the mean \pm s.e.m. For EcMutS–EcMutL complexes, kymographs of both channels were merged and dwell times were obtained from the overlapping regions of the kymographs. Since we never observed EcMutH DNA binding at physiological ionic conditions, the lifetime of the EcMutL–EcMutH complex was obtained from kymographs of EcMutH channel alone.

Binning method. All the binning methods in diffusion coefficient and lifetime analysis were described previously³⁶ and shown as below:

$$\begin{aligned} \text{Number of bins} &= [(\sqrt{N-1} - 1) \rightarrow \text{round up}] \\ \text{Size of bins} &= \left[\frac{\text{Max} - \text{min}}{\text{Number of bins} - 0.5} \right] \end{aligned}$$

where: N = numbers of events; max = maximum value of events; min = minimum value of events; bin start = min; bin end = min + (bin size \times number of bins). Bins distribution were then fit to standard curves³⁷.

EcMutS–EcMutL complex, EcMutL sliding clamp and EcMutL–EcMutH complex frequency measurements. EcMutS–EcMutL complex, EcMutS–EcMutL(R95F) complex, EcMutL sliding clamp, and EcMutL(R95F) sliding clamp frequency measurements were performed at a 600-ms frame rate with a 300-ms laser exposure time. Following the infusion of proteins, single-molecule movies (10 min) were recorded either immediately (Figs. 2h, 3g) or after 10 min incubation to establish equilibrium (Fig. 3d). EcMutL or EcMutL(R95F) were tracked by DiaTrack and the trajectories were obtained as described above. EcMutL or EcMutL(R95F) with a minimum lifetime of 30 s and a minimum diffusion coefficient of $0.1 \mu\text{m}^2 \text{s}^{-1}$ were counted as the number of EcMutL sliding clamps (N_L or $N_{L(R95F)}$). EcMutS and EcMutL (or EcMutL(R95F)) channels were merged and co-localized molecules with a minimum lifetime of 10 s were counted as EcMutS–EcMutL or EcMutS–EcMutL(R95F) complexes (N_{SL} or $N_{SL(R95F)}$). Following the real-time single-molecule recording, the number of DNA molecules (N_{DNA}) was determined by Sytox Orange staining. The frequencies of EcMutL sliding clamps (F_L), EcMutL(R95F) sliding clamps ($F_{L(R95F)}$), EcMutS–EcMutL complexes (F_{SL}) and EcMutS–EcMutL(R95F) complexes ($F_{SL(R95F)}$) were calculated using the following equations that also included corrections for labelling efficiencies of the proteins (the numbers in the denominator, Extended Data Table 1):

$$\begin{aligned} F_L &= \frac{N_L}{N_{DNA} \times 0.33} & F_{L(R95F)} &= \frac{N_{L(R95F)}}{N_{DNA} \times 0.29} \\ F_{SL} &= \frac{N_{SL}}{N_{DNA} \times 0.33 \times 0.45} & F_{SL(R95F)} &= \frac{N_{SL(R95F)}}{N_{DNA} \times 0.29 \times 0.45} \end{aligned}$$

EcMutL–EcMutH complex frequency measurement experiments were performed at a 600-ms frame rate with a 300-ms laser exposure time by using unlabelled

EcMutS (5 nM), EcMutL–Cy3 (10 nM) and EcMutH–AF647 (5 nM). Following the infusion of proteins, single-molecule movies (10 min) were recorded immediately (Fig. 4a). EcMutH was tracked using DiaTrack and the trajectories were obtained as described above. EcMutH particles with a minimum lifetime of 10 s were counted as the number of EcMutL–EcMutH complex (N_{LH}). Following the real-time single-molecule recording, the number of DNA molecules (N_{DNA}) was determined by Sytox 59 staining. The frequencies of EcMutL–EcMutH complex (F_{LH}) were calculated using the following equation that also included corrections for labelling efficiencies of the proteins (the numbers in the denominator, Extended Data Table 1):

$$F_{LH} = \frac{N_{LH}}{N_{DNA} \times 0.5}$$

All single-molecule studies were performed at least two separate times.

Analysis of the distribution of EcMutS–EcMutL complex dissociation processes. Following the infusion of proteins, single-molecule movies (10 min) were recorded after a 10 min incubation to establish equilibrium (Figs. 2a, 3e). Kymographs of EcMutS–EcMutL complex or EcMutS–EcMutL(R95F) complex were obtained by merging the two channels. We included only the initial events for individual molecules in the analysis when multiple events between the same molecules were observed.

MMR complementation *in vivo*. P. L. Foster (Indiana University) supplied the *E. coli* strains (all derivatives of MG1655 F- λ *ilvG-rfb-50 rph-1*). Mutation rates to rifampicin resistance were determined as previously described³⁴, using at least seven independent colonies for each genotype. ΔmutS , ΔmutL and ΔmutH strains were co-transformed with respective MMR protein expression plasmids (Extended Data Table 2) and pTARA plasmid³⁸ (for T7 RNA polymerase expression, a gift from K. Matthews, Addgene plasmid 31491). Single colonies were picked and grown overnight in the presence of $50 \mu\text{g ml}^{-1}$ kanamycin, $35 \mu\text{g ml}^{-1}$ chloramphenicol, 0.2% Arabinose, and IPTG (0.05 mM for EcMutS, 0.2 mM for EcMutL and 0.1 mM for EcMutH). As controls, single colonies of wild-type, ΔmutS , ΔmutL and ΔmutH strains without plasmids were similarly grown overnight. Dilutions of overnight cultures were plated on LB-Agar plates with or without $100 \mu\text{g ml}^{-1}$ rifampicin and allowed to grow overnight at 37°C. The colonies on LB with or without rifampicin were counted and the mutation rates were determined by fluctuation analysis³⁹.

EcMutH endonuclease activity assay. The d(GATC) endonuclease activity associated with EcMutH was assayed using 50 nM of the unmethylated FAM-labelled DNA substrate incubated with $4 \mu\text{M}$ EcMutH in a $50 \mu\text{l}$ reaction containing 20 mM Tris–HCl (pH 7.5), 5 mM MgCl_2 and 0.1 mM DTT at 37°C for 1 h. DNA was then denatured in 50% formamide at 95°C for 10 min and resolved on a 12% denaturing PAGE gel.

Analysis of EcMutS–EcMutL complex by smFRET. Particle positions were tracked with DiaTrack and particle trajectories were determined as previously described³⁶. The fluorescent intensities of the particles in both channels (I_{AF555} and I_{AF647}) were obtained from the trajectories using a custom written MATLAB script. A background correction was applied to all the intensities by using intensities in blank area. The apparent FRET efficiency was calculated as $I_{AF647}/(I_{AF555} + I_{AF647})$.

Average free-energy barrier calculation for the EcMutS–EcMutL complex. To determine whether EcMutS–EcMutL complex diffusion includes rotational as well as translation diffusion we calculate the free-energy barrier of rotation along the DNA in aqueous solution (ϵ)^{23,36}. These interactions can be described by diffusion on a rugged free-energy landscape by the following equation:

$$D = b^2 \frac{k_B T}{[6\pi\eta Rb^2 + 8\pi\eta R^3 + 6\pi\eta R(R_{OC})^2]} F(\epsilon)$$

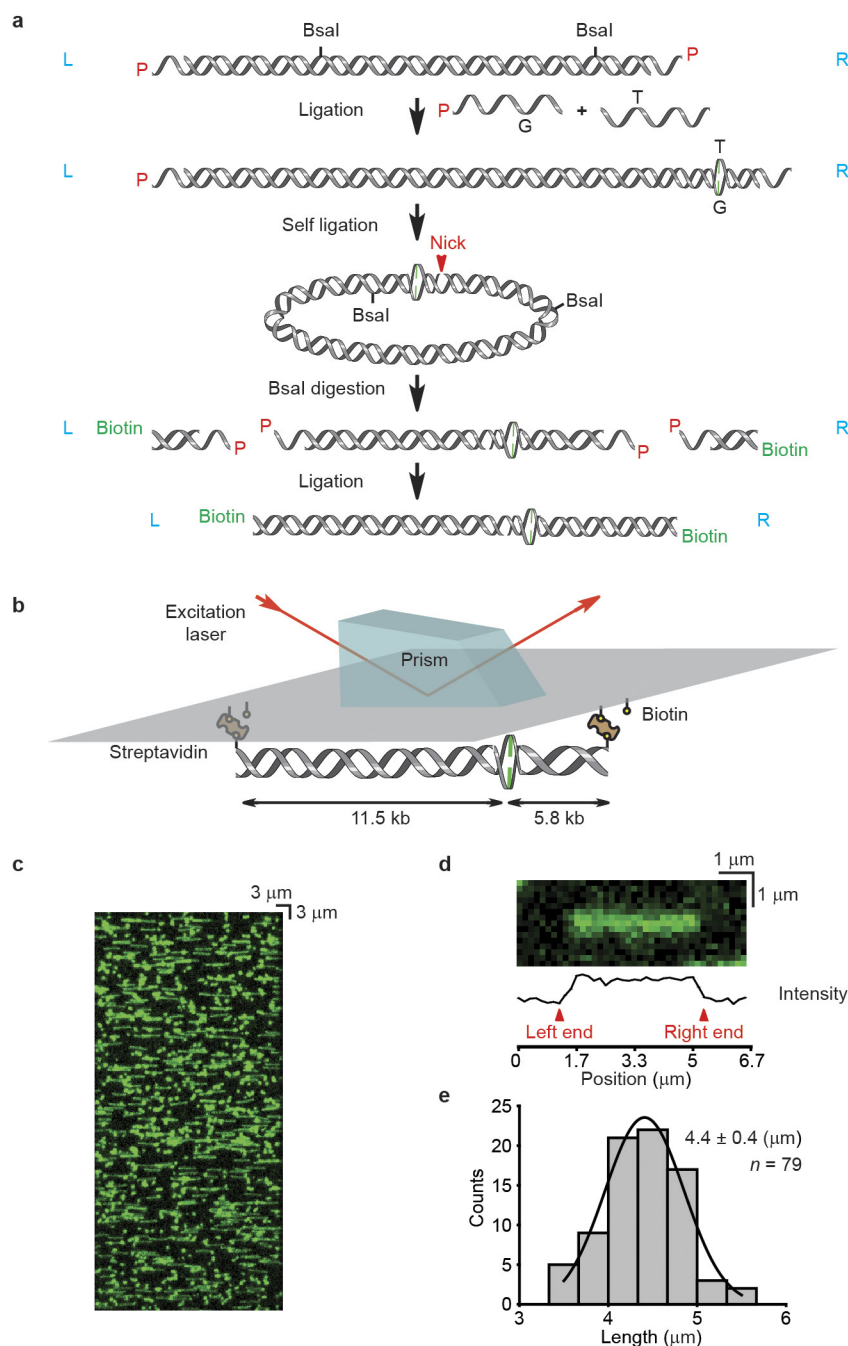
Where:

k_B = Boltzmann constant; T = temperature (Kelvin); η = viscosity of water; R = Stoke's radius of the complex; $b = (10B_d)/(2\pi)$, describes the effect of the helical pitch of DNA with regards to sliding; B_d = distance between two base pairs (0.34 nm); R_{OC} = minimum distance between the protein centre of mass and the DNA axis. Here, R_{OC} is estimated to be zero, because the centre of mass of MutS or MutL sliding clamp is close to the DNA axis; and $F(\epsilon) = \exp(-(\epsilon/k_B T)^2)$, describes the fluctuating part of the potential function that obeys a Gaussian distribution²³.

The Stoke's radii of an EcMutS dimer and an EcMutL dimer are 5 nm and 9 nm, respectively^{34,40}. There is no available Stoke's radius for the EcMutS–EcMutL complex; however, it is expected to be larger than that of EcMutL. Thus, we estimated a Stoke's radius of 9 nm for the EcMutS–EcMutL complex. The calculated ϵ in the equation using the experimentally determined diffusion coefficient for the EcMutS–EcMutL complex ($0.004 \pm 0.002 \mu\text{m}^2 \text{s}^{-1}$) yields an energy landscape of $1.7 \pm 0.2 k_B T$ (s.d.).

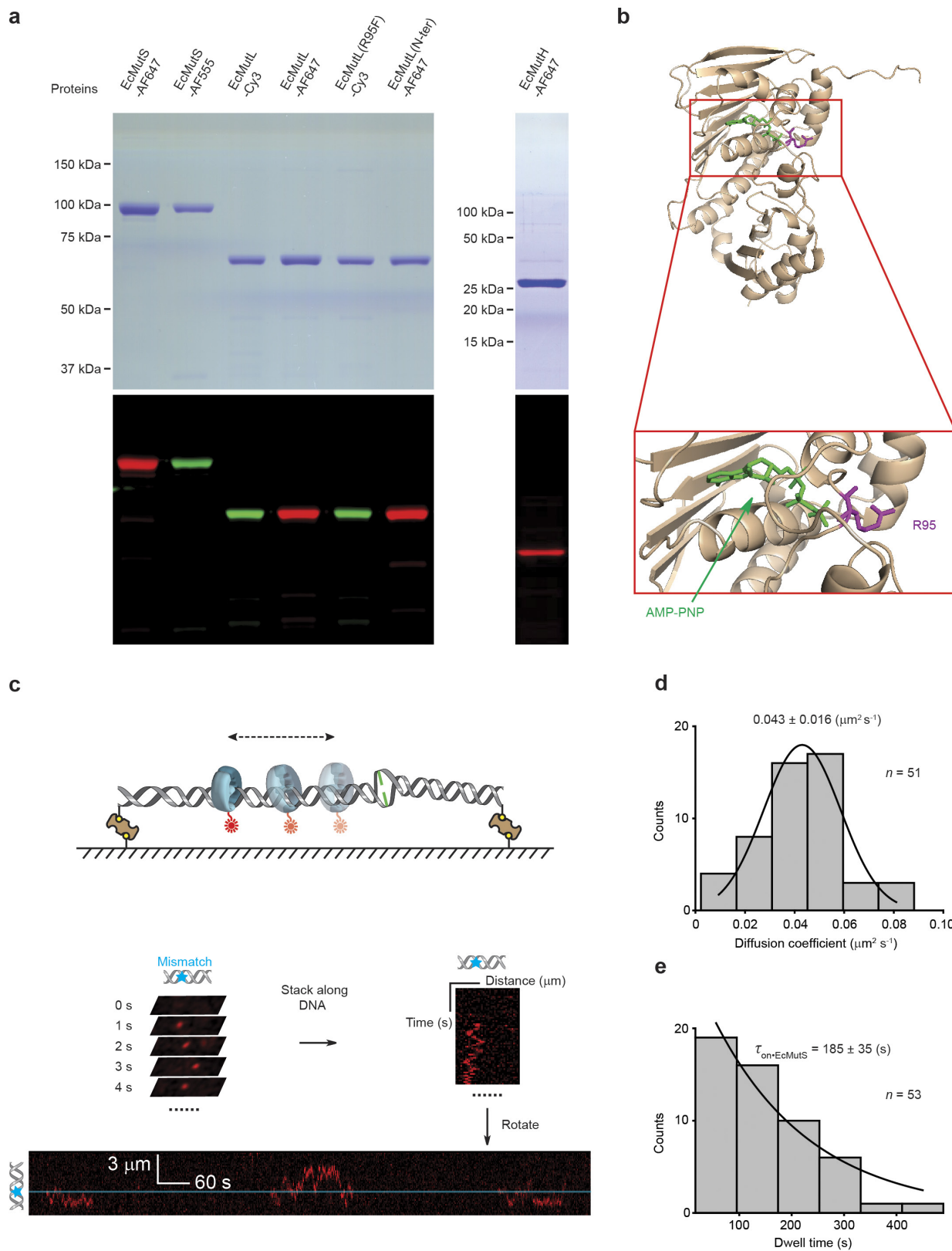
Data availability. The data that support the findings of this study are available from the authors upon reasonable request.

33. Senavirathne, G. *et al.* Widespread nuclease contamination in commonly used oxygen-scavenging systems. *Nat. Methods* **12**, 901–902 (2015).
34. Mendillo, M. L., Putnam, C. D. & Kolodner, R. D. *Escherichia coli* MutS tetramerization domain structure reveals that stable dimers but not tetramers are essential for DNA mismatch repair *in vivo*. *J. Biol. Chem.* **282**, 16345–16354 (2007).
35. Edelstein, A., Amodaj, N., Hoover, K., Vale, R. & Stuurman, N. Computer control of microscopes using microManager. *Curr. Protoc. Mol. Biol.* **Chapter 14**, Unit14.20 (2010).
36. Jones, N. D. *et al.* Retroviral intasomes search for a target DNA by 1D diffusion which rarely results in integration. *Nat. Commun.* **7**, 11409 (2016).
37. Qian, H. & Kou, S. C. Statistics and related topics in single-molecule biophysics. *Annu. Rev. Stat. Appl.* **1**, 465–492 (2014).
38. Wycuff, D. R. & Matthews, K. S. Generation of an AraC–araBAD promoter-regulated T7 expression system. *Anal. Biochem.* **277**, 67–73 (2000).
39. Lea, D. E. & Coulson, C. A. The distribution of the numbers of mutants in bacterial populations. *J. Genet.* **49**, 264–285 (1949).
40. Guarné, A. *et al.* Structure of the MutL C-terminal domain: a model of intact MutL and its roles in mismatch repair. *EMBO J.* **23**, 4134–4145 (2004).



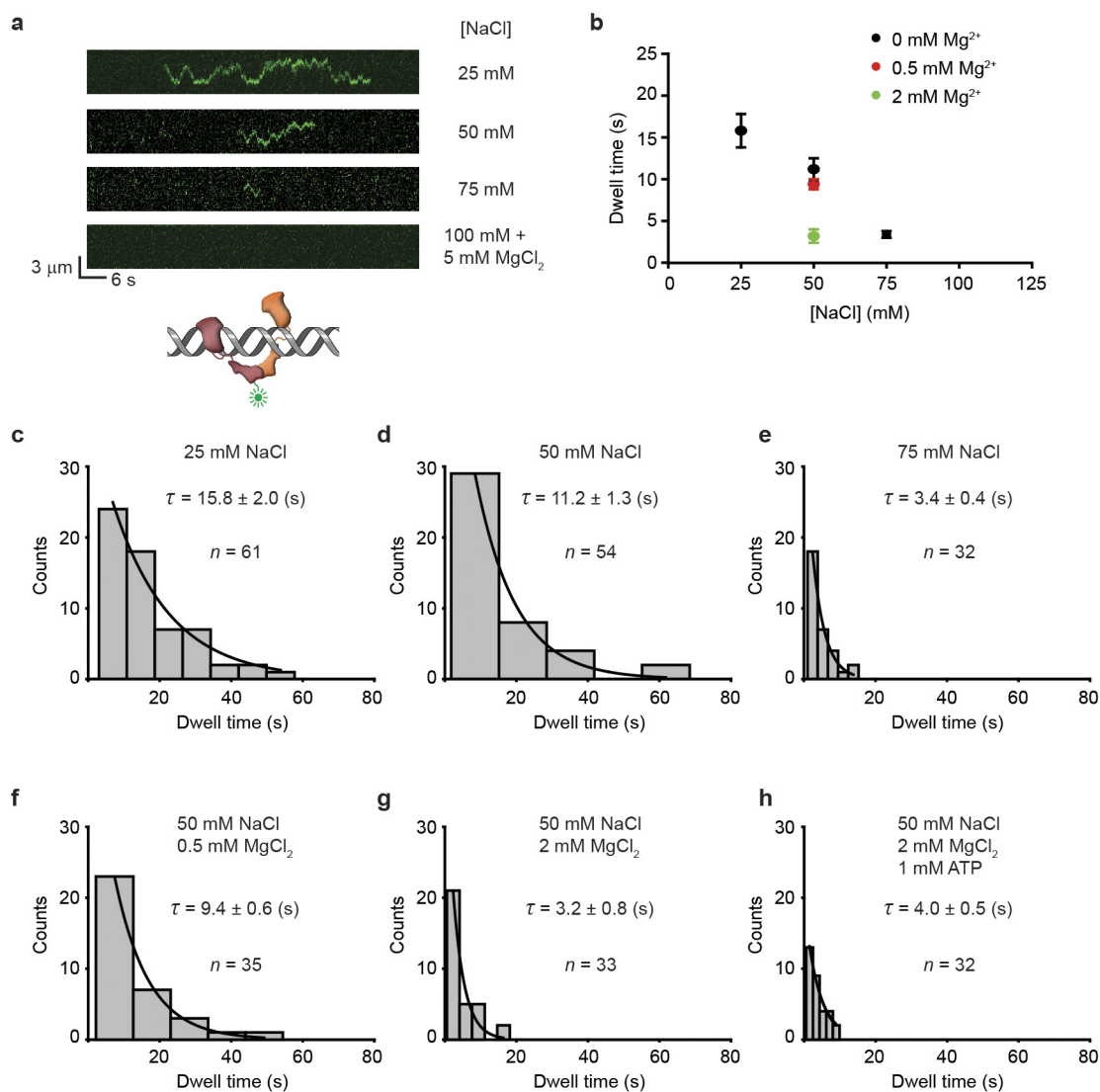
Extended Data Figure 1 | The construction of mismatched DNA used in single-molecule total internal reflection fluorescence (smTIRF) microscopy. **a**, A schematic illustration for the construction of a 17.3-kb mismatched DNA. L or R (blue) indicates the orientation of the DNA relative to the L and R cos end of λ -phage DNA. P (red) indicates the 5'-phosphate of the DNA. **b**, A schematic illustration of 17.3-kb mismatched DNA observation by prism-based smTIRF microscopy.

c, Representative mismatched DNA visualized by smTIRF microscopy in the absence of flow. The DNA was stained with Sytox Orange and a $40 \times 85 \mu\text{m}$ field of view is shown. **d**, A schematic illustration of the DNA length determination. **e**, The length distribution of the mismatched DNA observed by smTIRF microscopy. Gaussian fit of the data are shown along with the mean \pm s.d.

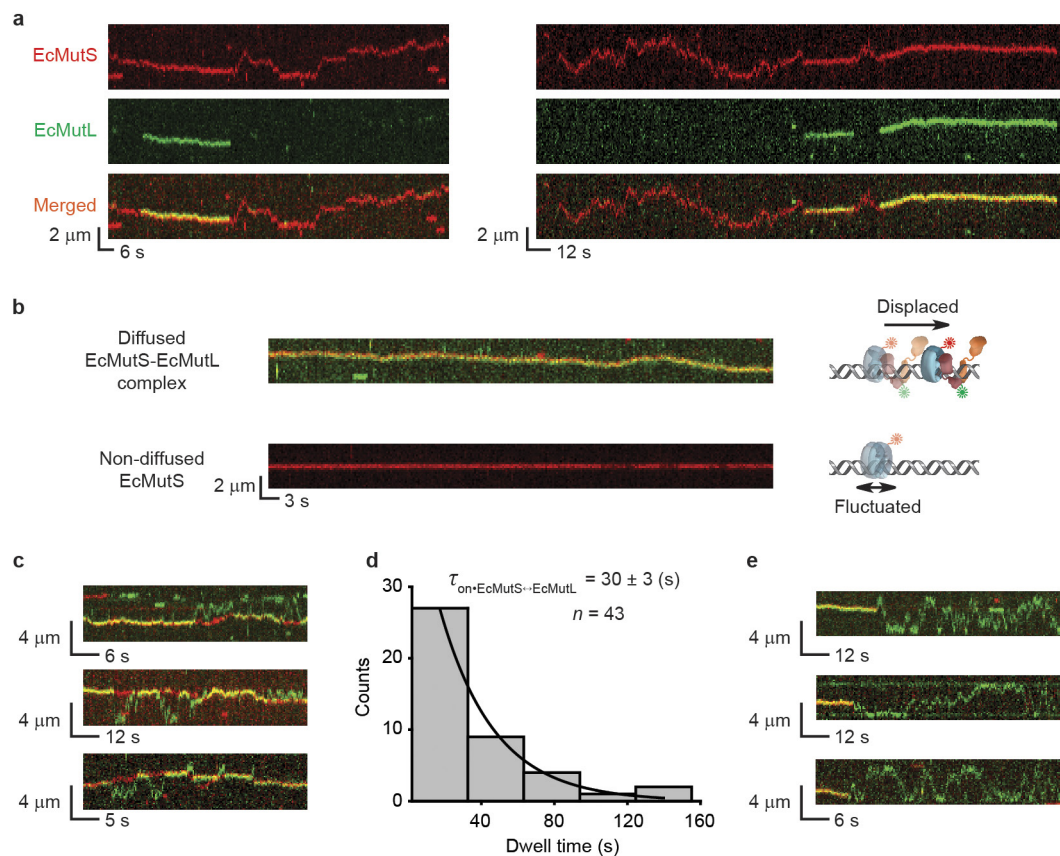


Extended Data Figure 2 | The fluorophore-labelled *E. coli* MMR proteins used in these studies and the formation of an EcMutS sliding clamp on DNA. a, Coomassie stained (top) and fluorescent (bottom) SDS-PAGE gels of labelled MMR proteins. For gel source data, see Supplementary Fig. 1. **b,** A crystal structure of the N-terminal domain of EcMutL bound to AMP-PNP (top) and magnification of the binding domain (bottom; PDB ID: 1B63). AMP-PNP is shown in green and

Arg-95 (R95) is shown in magenta²⁴. **c,** An illustration of the kymograph construction of three separate EcMutS sliding clamps on a single mismatched DNA. **d,** The distribution of diffusion coefficients for the EcMutS sliding clamp. The data were fit to a Gaussian with the mean \pm s.d. **e,** The distribution of dwell times (mean \pm s.e.m.) for the EcMutS sliding clamp.

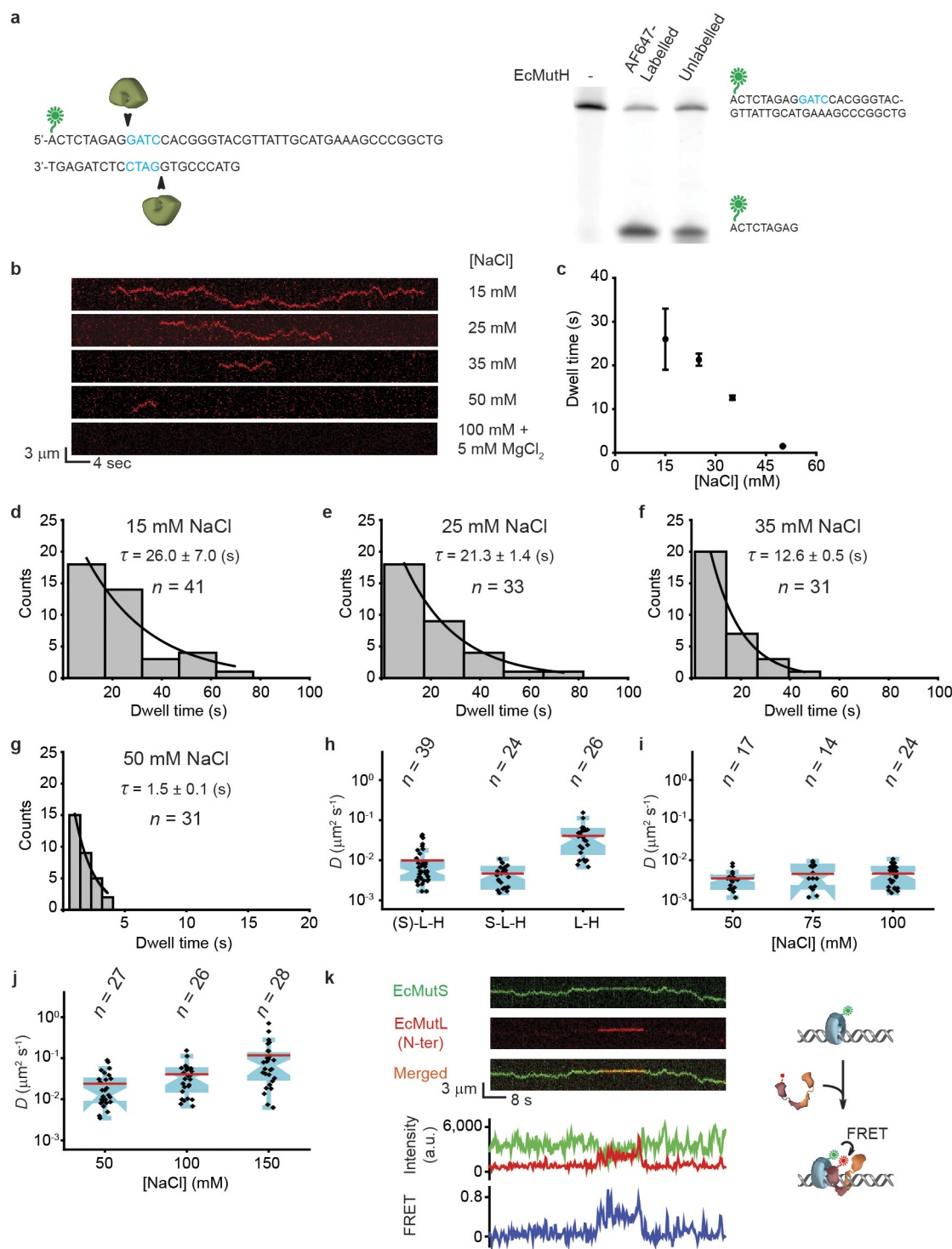


Extended Data Figure 3 | EcMutL does not bind DNA in physiological ionic conditions. **a, b**, Representative kymographs and dwell-times (mean \pm s.e.m.) for EcMutL binding to a mismatched DNA at various conditions. **c–h**, The distributions of dwell times (mean \pm s.e.m.) for EcMutL on mismatched DNA at different biochemical conditions as indicated.



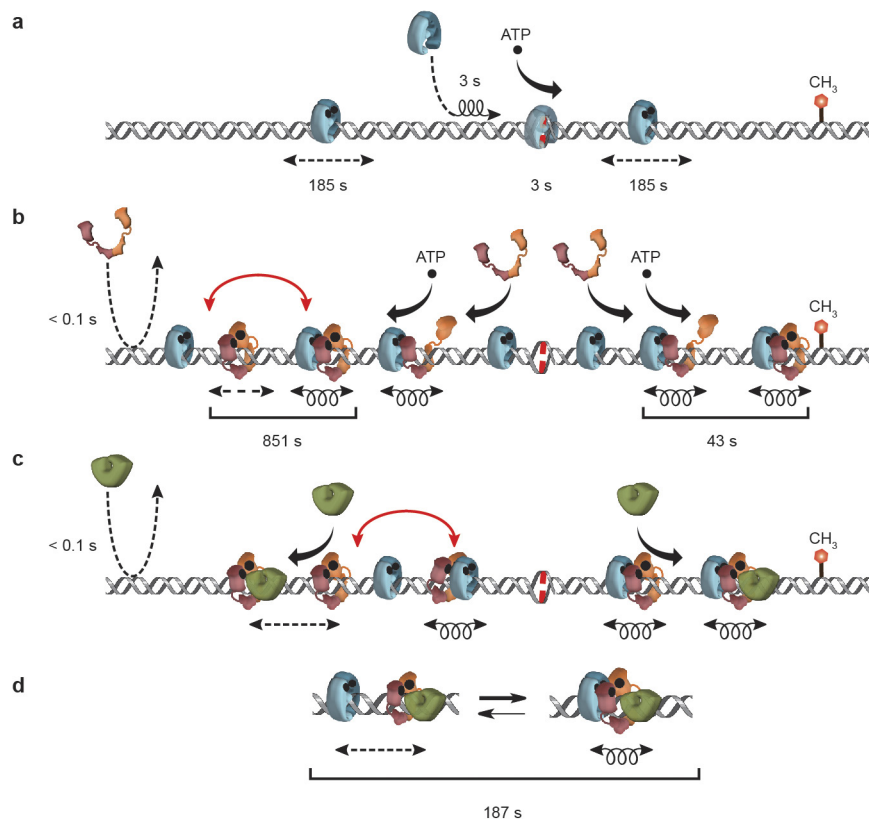
Extended Data Figure 4 | Representative kymographs of EcMutS–EcMutL complex and EcMutL particles. **a**, Representative kymographs showing the loading of EcMutL (green) on DNA by EcMutS sliding clamp (red). **b**, Representative kymographs of a diffusing EcMutS–EcMutL complex (merged channels) and a non-diffusing EcMutS (red) in the same field of view. The static kymograph of a non-diffusing EcMutS indicates that the change in protein position caused by microscope stage

drifting is negligible. **c**, Representative kymographs of oscillating EcMutS–EcMutL complexes. Two channels (red, EcMutS; green, EcMutL) were merged. **d**, The distribution of the association times (mean \pm s.e.m.) for EcMutS–EcMutL complexes during the oscillating phase. **e**, Representative kymographs of fast-diffusing EcMutL dissociation from EcMutS–EcMutL complexes. Two channels (red, EcMutS; green, EcMutL) were merged.



Extended Data Figure 5 | EcMutH lifetime on DNA and diffusion coefficient of EcMutS–EcMutL–EcMutH and/or EcMutL–EcMutH complex. **a**, A schematic illustration of EcMutH endonuclease assay (left) and a comparison of labelled or unlabelled EcMutH endonuclease activities (right). For gel source data, see Supplementary Fig. 1. **b**, **c**, Representative kymographs and dwell times (mean \pm s.e.m.) showing EcMutH on a single mismatched DNA under various ionic and magnesium conditions. **d**–**g**, The distributions of dwell times (mean \pm s.e.m.) for EcMutH on a single mismatched DNA at different biochemical conditions as indicated. **h**, Box plots showing D for oscillating (EcMutS)–EcMutL–EcMutHAF647 ((S)–L–H) complex; the established EcMutSAF555–EcMutL–EcMutHAF647 complex (S–L–H); and free EcMutL–EcMutHAF647 complex (L–H) at 100 mM NaCl. **i**, Box plots showing D for established EcMutS–EcMutL–EcMutH complex

at different NaCl concentrations. Two-sample t -test showed no significant difference between diffusion coefficients ($P > 0.1$). **j**, Box plots showing D for free EcMutL–EcMutH complex at different NaCl concentrations. Two-sample t -test showed significant differences between diffusion coefficients ($P < 0.05$). **k**, Top left, representative kymographs showing FRET between C-terminal AF555-labelled EcMutS and N-terminal AF647-labelled EcMutL (N-ter). Bottom, fluorescent intensities of EcMutS–AF555 (donor, green), EcMutL–AF647 (acceptor, red) and FRET (blue) between them when only the green laser was used for illumination. Right, a schematic illustration of kymographs. Experimental FRET measure ($E_{\text{EcMutS–EcMutL}} = 0.48 \pm 0.05$; mean \pm s.d.) and theoretical FRET ($E_{\text{EcMutS–EcMutL}} = 0.56$) based on crosslink structure²² appeared comparable. n = number of molecules throughout.



Extended Data Figure 6 | The interactions and kinetic properties of the molecular switch/sliding clamp mechanism for *E. coli* MMR.

a, Illustration of the kinetics and diffusion properties of EcMutS. **b**, Illustration of the kinetics and diffusion of EcMutS with EcMutL. **c**, Illustration of the kinetics and diffusion of EcMutS, EcMutL and EcMutH. **d**, Oscillation dynamics of the EcMutS–EcMutL–

EcMutH complex (see main text). Coil, 1D-diffusion search along the backbone; dashed straight arrow, rotation-independent 1D-diffusion; black curved arrow, binding; red curved arrow, oscillating complex; dashed curved arrow, binding-dissociation; binding times and ATP (•) are indicated.

Extended Data Table 1 | The oligonucleotides and fluorophore-labelled MMR proteins used in these studies

Name*	Sequences
EcMutL(R95F) for	GCCATTATCAGCCTGGGCTTTTTCGGTGAGGCGCTGGCGAGTATC
EcMutL(R95F) rev	GATACTCGCCAGCGCCTCACCGAAAAAGCCCAGGCTGATAATGGC
EcMutL(N-ter) for	GGAGGCCATATGCTCTGCACACCGTCGCGTGAGGGGCATCACCATCACCATCATTCAATGCCAA TTCAGGTCTTACC
EcMutL(N-ter) rev	TCCGCCCTCGAGTTACTCATCTTTCAGGGCTTTTATCG
Lambda mismatch 1	Phos-GGGCGGCGACCTGCTTAGGATCATCGAGGATCGACGTCGGTGCAATTCAGCGGACTAGTCC
Lambda mismatch 2	AGGTCGCCGCCCCGACTAGTCCGCTGAATTGCACCGACGTTGATCCTCGATGATCCTAAGC
Lambda linker 1	Phos-TTCTTGAGTCCCCTGCAGCGATTAATACGACTAGAGC
Lambda linker 2	Biotin-GCTCTAGTCGTATTAATCGCTGCAGGGGACTCCA
Lambda linker 3	Phos-CGCTTAGTGCTATGATGCGTTCGATCACTCCATGTGATC
Lambda linker 4	Biotin-GATCACATGGAGTGATCGAACGCATCATAGCACTA
EcMutH for1	AGGGCATCACCATCACCATCATTCAATGTCCCAACCTCGCCCA
EcMutH for2	GGAGGCCATATGCTCTGCACACCGTCGCGTGAGGGGCATCACCATACCA
EcMutH rev	GGAGGCGGATCCCTACTGGATCAGAAAATGAC

Protein†	Labelled monomer	Unlabelled dimer	Dimer with a single fluorophore	Dimer with two fluorophores
AF647 labelled EcMutS	26%	55%	38%	7%
Cy3 labelled EcMutL	18%	67%	30%	3%
Cy3 labelled EcMutL(R95F)	16%	71%	26%	3%
AF647 labelled EcMutH	50%	N/A	N/A	N/A

*Oligonucleotides used in these studies.

†Fluorophore-labelling efficiencies of MMR proteins used in these studies.

Extended Data Table 2 | Cellular complementation and mutation rates of MMR proteins

Strain	MMR proteins in expression plasmids*	Genotype	Mutation rates [†] (per 10 ⁹ cells)	Range (per 10 ⁹ cells)
MG 1655	-	wild type	5.2	1.8-9.0
MG 1655 $\Delta mutS$	-	<i>mutS</i> ⁻	522.0	334.9-860.8
MG 1655 $\Delta mutS$	EcMutS (D835R, R840E)-his ₆ -ald ₆	<i>MutS</i> -(D835R, R840E)	1.2	0.4-12.9
MG 1655 $\Delta mutL$	-	<i>mutL</i> ⁻	1004.9	572.2-3481.1
MG 1655 $\Delta mutL$	EcMutL-his ₆ -ald ₆	<i>MutL</i> ⁺	6.3	0.4-9.5
MG 1655 $\Delta mutL$	EcMutL (R95F)-his ₆ -ald ₆	<i>MutL</i> (R95F)	702.3	29.5-1981.3
MG 1655 $\Delta mutL$	ald ₆ -his ₆ -EcMutL	<i>MutL</i> ⁺	0.8	0.4-1.2
MG 1655 $\Delta mutH$	-	<i>mutH</i> ⁻	100.2	77.8-371.9
MG 1655 $\Delta mutH$	ald ₆ -his ₆ -EcMutH	<i>MutH</i> ⁺	3.3	1.2-5.7

*Amino acid substitutions are indicated in parenthesis; the arrangement of the hexa-histidine (his₆) and formylglycine-generating enzyme (ald₆) recognition sequence relative to the N terminus (beginning) or C terminus (end) of the MMR gene.

[†]The median mutation rates are reported.

Extended Data Table 3 | The frequency of particle varieties observed by single-molecule analysis

Proteins	Diffusion	Co-localize with stained DNA	Ratio		Number of molecules
EcMutS only	+	+	88%	93%	56
	+	-*	5%		
	-†	+	7%		
EcMutS-EcMutL complex	+	+	69%	80%	54
	+	-*	11%		
	-	+	4%		
	-	-	16%		
EcMutS-EcMutL (R95F) complex	+	+	76%	85%	46
	+	-*	9%		
	-	+	4%		
	-	-	11%		
EcMutL-EcMutH complex	+	+	80%	89%	95
	+	-*	9%		
	-	+	1%		
	-	-	10%		

*The DNA was probably broken during staining.

†Immobile particles bound near the mismatch.

Extended Data Table 4 | Diffusion coefficients

Protein conformation	[NaCl] (mM)	Mean \pm S.D. ($10^{-3} \mu\text{m}^2 \text{s}^{-1}$)
EcMutS sliding clamp	100	43 \pm 16
EcMutS-EcMutL complex	50	4 \pm 2
	75	4 \pm 3
	100	4 \pm 2
	125	5 \pm 3
	150	5 \pm 3
Fast diffusing EcMutL	25	453 \pm 311
	100	888 \pm 393
	150	1716 \pm 623
	200	2220 \pm 778
EcMutS-EcMutL(R95F) complex	100	3 \pm 2
(EcMutS)-EcMutL-EcMutH*	100	10 \pm 10 [†]
EcMutS-EcMutL-EcMutH [‡]	50	4 \pm 2
	75	5 \pm 3
	100	5 \pm 3
Free EcMutL-EcMutH complex [§]	50	24 \pm 23
	100	41 \pm 34
	150	118 \pm 151

*Oscillating complexes potentially containing unlabelled EcMutS (EcMutS) with co-localized EcMutL-Cy3 and EcMutH-AF647.

[†]A two-sample *t*-test suggests a significant difference ($P = 0.001$) in diffusion coefficient between the EcMutS-EcMutL complex and (EcMutS)-EcMutL-EcMutH at 100 mM NaCl.

[‡]Complexes containing co-localized EcMutS-AF555, unlabelled EcMutL and EcMutH-AF647.

[§]See Methods.

^{||}A two-sample *t*-test suggests a significant difference ($P < 0.0001$) in diffusion coefficient between the free EcMutL-EcMutH complex and (EcMutS)-EcMutL-EcMutH at 100 mM NaCl.

Genetic and mechanistic diversity of piRNA 3'-end formation

Rippe Hayashi^{1*}, Jakob Schnabl^{1*}, Dominik Handler¹, Fabio Mohn[†], Stefan L. Ameres¹ & Julius Brennecke¹

Small regulatory RNAs guide Argonaute (Ago) proteins in a sequence-specific manner to their targets and therefore have important roles in eukaryotic gene silencing¹. Of the three small RNA classes, microRNAs and short interfering RNAs are processed from double-stranded precursors into defined 21- to 23-mers by Dicer, an endoribonuclease with intrinsic ruler function. PIWI-interacting RNAs (piRNAs)—the 22–30-nt-long guides for PIWI-clade Ago proteins that silence transposons in animal gonads—are generated independently of Dicer from single-stranded precursors^{2,3}. piRNA 5' ends are defined either by Zucchini, the *Drosophila* homologue of mitoPLD—a mitochondria-anchored endonuclease^{4,5}, or by piRNA-guided target cleavage^{6,7}. Formation of piRNA 3' ends is poorly understood. Here we report that two genetically and mechanistically distinct pathways generate piRNA 3' ends in *Drosophila*. The initiating nucleases are either Zucchini or the PIWI-clade proteins Aubergine (Aub) or Ago3. While Zucchini-mediated cleavages directly define mature piRNA 3' ends^{8,9}, Aub/Ago3-mediated cleavages liberate pre-piRNAs that require extensive resection by the 3'-to-5' exoribonuclease Nibbler (*Drosophila* homologue of Mut-7)^{10–13}. The relative activity of these two pathways dictates the extent to which piRNAs are directed to cytoplasmic or nuclear PIWI-clade proteins and thereby sets the balance between post-transcriptional and transcriptional silencing. Notably, loss of both Zucchini and Nibbler reveals a minimal, Argonaute-driven small RNA biogenesis pathway in which piRNA 5' and 3' ends are directly produced by closely spaced Aub/Ago3-mediated cleavage events. Our data reveal a coherent model for piRNA biogenesis, and should aid the mechanistic dissection of the processes that govern piRNA 3'-end formation.

piRNA biogenesis is initiated by endonucleolytic definition of piRNA 5' ends (Fig. 1a). Based on the nuclease involved, this defines primary (processed by Zucchini)^{4,5} and secondary (processed by Aub/Ago3)^{6,7} piRNAs. Attributes of these two endo-cleavages are a 5' uridine signature for primary piRNAs, and a 10-nt 5' sense/antisense offset for secondary piRNAs (ping-pong signature). Following 5'-end cleavage, piRNA intermediates are anchored with their 5' ends in PIWI proteins before their 3' ends are matured^{14,15} (we refer to these as pre-piRNAs). Zucchini also liberates 3' ends of primary and secondary piRNAs^{8,9}. In *zucchini* mutants, however, Aub/Ago3-bound secondary piRNAs are still abundant¹⁶, indicative of alternative 3'-end formation pathways (Extended Data Fig. 1a). The 3' ends of these piRNAs lack signs for endonucleolytic processing, such as a coupling signature stemming from Zucchini-mediated, phased piRNA biogenesis^{8,9}, or a 3'/5' ping-pong signature indicative of slicer-mediated 3'-end formation (Extended Data Fig. 1b). This supports the hypothesis that Zucchini-independent piRNA 3' biogenesis involves exonucleolytic resection of pre-piRNAs that have been generated by Aub/Ago3 (Fig. 1a)⁹.

We identified piRNA 5' species (piRNAs with the same 5' end) that—besides piRNAs in the 23–29 nt range—exhibit one abnormally

long isoform that extends to the cleavage position of a complementary piRNA (Fig. 1b). These isoforms are also found in libraries from immuno-purified PIWI proteins (Fig. 1b), indicating that they represent Aub/Ago3-loaded pre-piRNAs whose 3' ends have been formed by slicing and await trimming. Consistent with this, the long isoforms lack 2'-O-methylation at their 3' ends (Fig. 1b).

To identify the 3' exonuclease involved, we used a piRNA biogenesis reporter that recapitulates Zucchini-independent piRNA 3'-end formation⁹; expression of a reporter with two target sites for cellular piRNAs forces the generation of responder piRNAs in Zucchini-depleted ovaries (Fig. 1c, d). We combined this reporter with a double-shRNA expression cassette to co-deplete Zucchini and a gene of interest (Extended Data Fig. 2a–c). A strong candidate for the exonuclease is the PARN-like nuclease PNLDC1, which trims pre-piRNAs in silkworm¹⁷. As PARN-family nucleases are absent in *Drosophila*, we tested instead the mitochondria-anchored Tudor/KH-domain protein Papi, an essential PNLDC1 co-factor in silkworm (Extended Data Fig. 2d, e)^{17,18}. Co-depletion of Zucchini and Papi does not impair piRNA generation from two independent reporters (Fig. 1d, Extended Data Fig. 2b, f), and global piRNA levels are comparable between Zucchini- versus Zucchini and Papi co-depleted ovaries (Extended Data Fig. 2g). Notably, Piwi-bound piRNAs increase by approximately 0.5 nt in length in *papi* mutants (Fig. 1e, Extended Data Fig. 3a, j)⁸. As 3' ends of Piwi-bound piRNAs are generated predominantly by Zucchini^{8,9}, we conclude that Papi-assisted piRNA trimming—if conserved in flies—occurs downstream of Zucchini, consistent with its role in mouse and silkworm^{8,9,17,18}.

We next tested the 3'-to-5' exoribonuclease Nibbler/Mut-7, which trims some micro RNAs (miRNAs) after their loading into Ago1^{10,11}, and which has been reported to modulate piRNA lengths^{12,13}. Co-depletion of Zucchini and Nibbler (Extended Data Fig. 2c) ablates piRNA production from both reporters despite trigger piRNAs remaining abundant and silencing-competent (Fig. 1d, Extended Data Fig. 2f, h).

Consistent with Nibbler acting on slicer-generated pre-piRNAs, it is enriched in perinuclear nuage together with Aub/Ago3, while Papi co-localizes with Zucchini at mitochondria (Fig. 1f, Extended Data Fig. 2d, e). In *aubergine* mutants, Nibbler's nuage localization is reduced, yet Nibbler does not enrich in Krimper foci where unloaded Ago3 accumulates (Fig. 1f)^{16,19,20}. Nibbler's co-localization with Aub/Ago3 therefore probably depends on these factors being loaded with pre-piRNAs. We did not detect robust interactions between Nibbler and Aub/Ago3 by co-immunoprecipitation (weak interactions between Nibbler and Piwi were detected¹³), hinting at a transient interaction (Extended Data Fig. 2i, j).

To characterize Nibbler's role in piRNA biogenesis we generated flies that express no detectable Nibbler protein (Extended Data Fig. 3b, c). As reported^{12,13}, *nibbler* mutants are viable and fertile, but defective in *mir-34* maturation (Extended Data Fig. 3d). Also as reported, localization and abundance of PIWI proteins, overall piRNA levels, and transposon silencing are not affected (Extended Data Fig. 3e–h). Average

¹Institute of Molecular Biotechnology of the Austrian Academy of Sciences (IMBA), Vienna Biocenter (VBC), Dr. Bohrgasse 3, 1030 Vienna, Austria. [†]Present address: Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland.

*These authors contributed equally to this work.

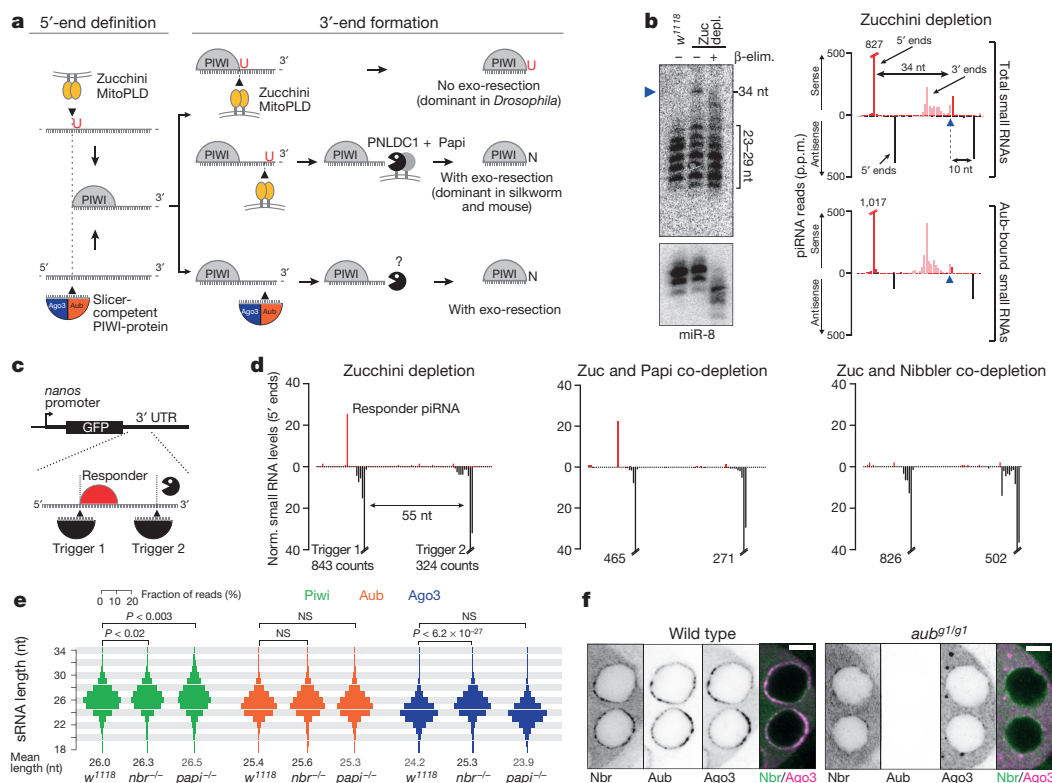


Figure 1 | The 3'-to-5' exonuclease Nibbler matures piRNA 3' ends from slicer-cleaved pre-piRNAs. **a**, Schematic illustration of piRNA 5' and 3' biogenesis. **b**, Northern blot against individual piRNA 5' species (mature piRNAs: 23–29 nt) detects a 34-nt pre-piRNA (blue arrowhead). 3'-end methylation probed by β -elimination (miR-8 serves as control). To the right, sequencing counts of the corresponding piRNAs (normalized to 1 million miRNA reads; p.p.m.) from total small RNAs or from an Aub immunoprecipitation are shown. **c**, Schematic of the dual-site piRNA

piRNA length, however, is mildly increased (Extended Data Fig. 3i; our sequencing libraries span 18–40 nt, Extended Data Fig. 4). Notably, this originates primarily from Ago3-bound piRNAs, which increase > 1 nt in length (Fig. 1e). This supports a specific role for Nibbler in resecting Aub/Ago3-generated pre-piRNAs. Indeed, somatic Piwi-bound piRNAs, whose 3' ends are generated by Zucchini, show no length change in *nibbler* mutants, in contrast to *papi* mutants (Extended Data Fig. 3j)^{12,13}. These results indicate that Nibbler does not fine-tune piRNA length as proposed^{12,13}, but instead represents the central exonuclease of a distinct piRNA 3'-end pathway that resects slicer-generated pre-piRNAs to mature piRNAs.

If Zucchini endonuclease and Nibbler exonuclease act in separate pathways to generate piRNA populations with similar overall length, the 3' profiles of piRNA 5' species should differ in single-mutant ovaries. We inspected individual Aub/Ago3-bound piRNA 5' species by northern blot analysis and sequencing (Fig. 2a, b). While piRNAs in Zucchini-depleted ovaries display a broad length profile (consistent with exonucleolytic resection), piRNAs in *nibbler* mutants display discrete length patterns with major isoforms typically being followed by uridine (downstream-U signature), a hallmark of Zucchini cleavages (Fig. 2a, b)^{8,9}.

To generalize these findings, we determined the downstream-U signature and the 3'-end precision index for thousands of piRNA 5' species bound to Piwi/Aub/Ago3. This allows several conclusions. (1) In agreement with the two-pathway model, the downstream-U signature increases in *nibbler* mutant ovaries, yet is ablated in Zucchini-depleted ovaries (Fig. 2c; Piwi- piRNAs are lost in the absence of Zucchini). (2) In wild-type ovaries, the downstream-U signature is strong for Piwi-bound piRNAs, intermediate for Aub, and very weak for Ago3 (Fig. 2c), indicating that Zucchini acts predominantly on Piwi and Aub.

biogenesis reporter. UTR, untranslated region. **d**, Levels of small RNAs (p.p.m.) mapping to the biogenesis reporter (5' ends only) in indicated genetic backgrounds. **e**, Length profiles of TE-mapping small RNAs isolated from Piwi/Aub/Ago3-immuno-precipitates (genetic background indicated; *P*-values calculated with two-sided *t*-test). **f**, Confocal images showing localization of GFP-Nibbler, Aub and Ago3 in *w¹¹¹⁸* or *aubergine* mutant egg chambers (scale bars, 5 μ m; individual channels as inverted grey scale images).

(3) The downstream-U signature correlates with the 3'-end precision index of piRNA populations (Fig. 2c). (4) In *nibbler* mutants, the 3'-end precision index increases for Piwi-, Aub-, and Ago3-bound piRNAs, indicating that Nibbler acts on all three PIWI proteins (Extended Data Fig. 5a). (5) Characteristics of piRNA 5' ends do not correlate with the 3'-end precision index (Extended Data Fig. 5b), arguing that 5'-end generation does not dictate the mode of 3'-end formation.

In agreement with Nibbler and Zucchini acting in parallel pathways, the length profiles of wild-type piRNAs appear to be a composite of the two respective single-pathway profiles (Fig. 2a, b). Inspired by this observation, we computed the relative contribution of Nibbler and Zucchini for Aub/Ago3-bound piRNAs. For each of the approximately 300 analysable piRNA 5' species, we determined the Zucchini/Nibbler contribution at which the combined length profile best mimics the wild-type profile (Fig. 2d). For both, Aub- and Ago3-bound species the wild-type profiles can be accurately modelled from the single-pathway profiles (Extended Data Fig. 5c). This results in an approximate median 70:30 dominance of Zucchini over Nibbler for Aub-bound piRNAs and an opposite ratio for Ago3-bound piRNAs (Fig. 2e), in agreement with the 3'-end characteristics of the respective piRNA populations in wild-type ovaries (Fig. 2c). Our data demonstrate that two parallel pathways with varying contributions form 3' ends of Aub/Ago3-bound piRNAs: Zucchini generates most Aub-bound piRNAs, while Nibbler generates most Ago3-bound piRNAs.

Zucchini-mediated piRNA 3'-end formation results in processing of the downstream precursor RNA into phased piRNAs bound to Piwi (referred to as triggering)^{8,9}. 3'-end formation by Nibbler instead prevents triggering owing to degradation of the downstream precursor (Fig. 3a). As Zucchini compensates for 3'-end formation in the absence of Nibbler, triggering levels should increase in *nibbler* mutants. Indeed,

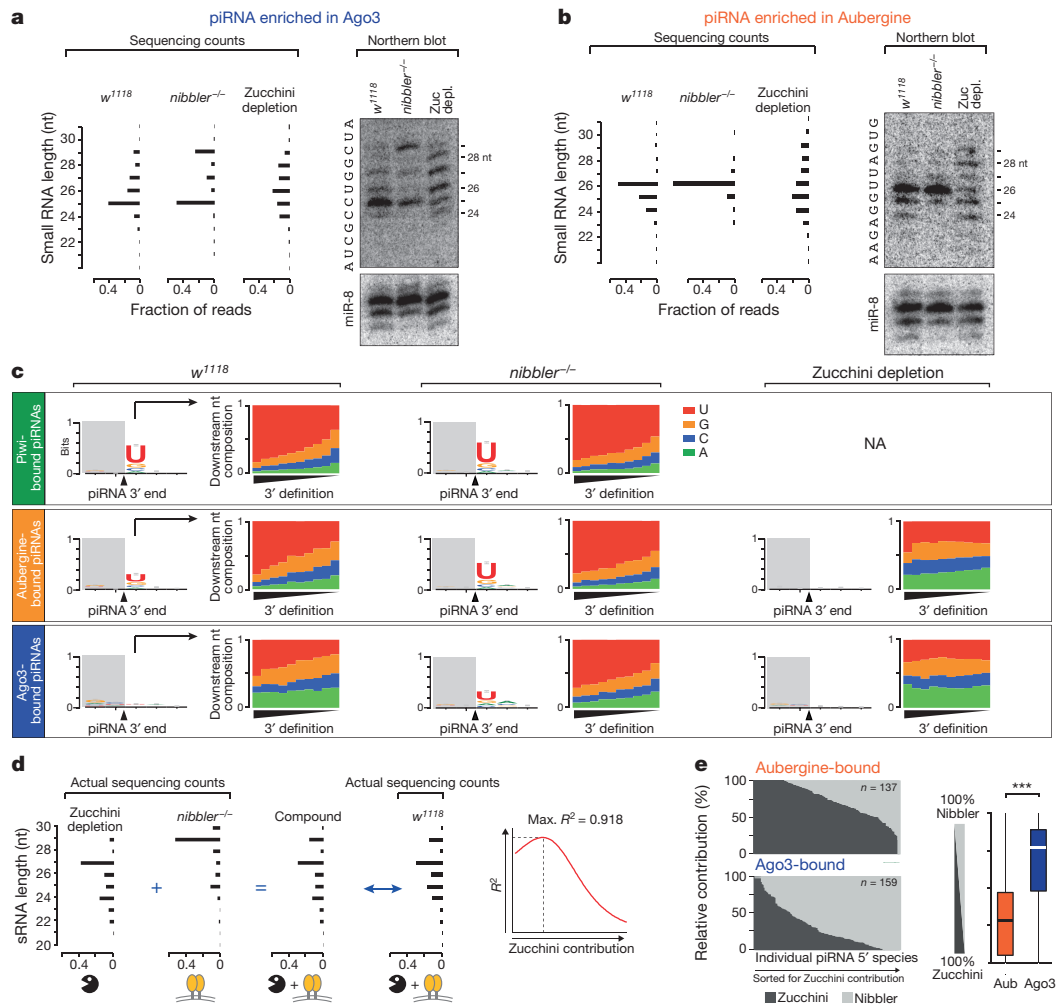


Figure 2 | Two genetically independent pathways generate piRNA 3' ends. **a, b**, Northern blot analysis (loading control: miR-8) and sequencing counts for Ago3- (**a**) or Aub- (**b**) enriched piRNAs in indicated genetic backgrounds. Bar charts display fraction of reads with indicated length. Respective RNA sequences are shown. **c**, Sequence logos display nucleotide bias around dominant 3' ends of Piwi/Aub/Ago3-bound piRNAs isolated from ovaries of indicated genetic background. Bar plots

the occurrence of Piwi-bound piRNA 5' ends immediately downstream of Aub/Ago3-bound piRNA 3' ends increases in *nibbler* mutant compared to wild-type ovaries (Fig. 3b). As expected, this increase is more pronounced for Ago3/Piwi linkages compared to Aub/Piwi linkages. We conclude that in wild-type ovaries, downstream slicing and subsequent Nibbler-catalysed pre-piRNA resection limits the extent of triggering, especially for Ago3-bound pre-piRNAs.

In agreement with elevated triggering, the levels of Piwi-bound piRNAs increase at the expense of ping-pong piRNAs in *nibbler* mutants (Fig. 3c). It is unclear why Aub-bound, but not Ago3-bound piRNAs are reduced in the absence of Nibbler. Possibly this is due to Ago3 incorporating abnormally high levels of antisense piRNAs (Extended Data Fig. 6a, b). As a consequence of the shifts in piRNA populations, ping-pong signatures for nearly all transposable elements (TEs) decrease (Extended Data Fig. 6c). This indicates a competitive relationship between two piRNA 3'-end formation pathways with consequences for piRNA loading into PIWI proteins: Nibbler limits the extent at which slicer-induced piRNA biogenesis propagates into Zucchini-mediated downstream piRNA biogenesis that fuels nuclear Piwi. By contrast, Zucchini consumes piRNA precursors, reducing their participation in ping-pong during post-transcriptional regulation.

These findings prompted us to re-examine the long-standing question of why secondary piRNA populations from some TEs remain

display the nucleotide composition at the position following dominant 3' ends (piRNAs are split into ten equally sized bins sorted for their precision index). **d**, Shown is how the Zucchini/Nibbler contribution to 3'-end formation of individual piRNA 5' species was determined. **e**, Shown are the Zucchini/Nibbler contributions for 3'-end formation for Aub- and Ago3-bound piRNAs (left, as stacked bar diagram; right, as box plot; Tukey definition; *** $P < 0.001$ after Wilcoxon rank-sum test).

abundant in Zucchini-depleted ovaries ('robust TEs'; for example, *Doc*), while others collapse ('sensitive TEs'; for example, *I*-element; Extended Data Fig. 6d)¹⁶. We reasoned that only TEs with a minimal abundance of Nibbler substrates (reflected by abundant ping-pong piRNAs) could maintain piRNA biogenesis in the absence of Zucchini. That is because Zucchini generates piRNA 3' ends independent of additional precursor cleavages, while Nibbler-mediated 3'-end formation requires a second piRNA-guided cleavage event close by (Fig. 3a). Indeed, ping-pong piRNA levels are substantially higher for robust compared to sensitive TEs in wild-type ovaries (Extended Data Fig. 6d). TEs with ping-pong piRNAs below a threshold level therefore cannot compensate for Zucchini loss, as the production of Nibbler substrates is too inefficient. As ping-pong is a feed-forward loop, this results in the collapse of piRNA biogenesis.

As the most direct test for two separate piRNA 3'-end pathways, we co-depleted Zucchini and Nibbler. As expected from the piRNA biogenesis reporter experiments (Fig. 1d), this results in piRNA loss for some TEs, which generate abundant piRNAs in Zucchini-depleted ovaries (Extended Data Figs 1a and 7a). Surprisingly, however, piRNAs that map to several other TEs are only mildly affected (Extended Data Fig. 7a, b), total germline piRNA levels are reduced less than twofold compared to Zucchini-depleted ovaries (Extended Data Fig. 7c), and TE derepression is similar to Zucchini-depleted ovaries (Extended Data

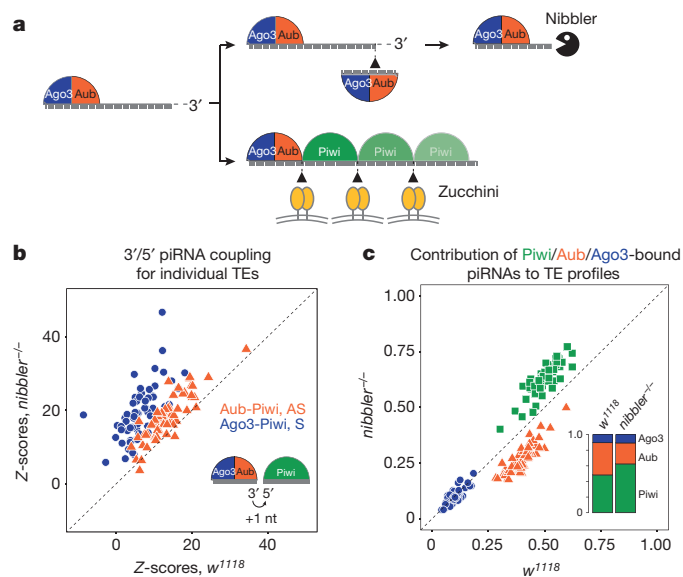


Figure 3 | Nibbler and Zucchini set the balance between primary and secondary piRNA biogenesis. **a**, Schematic of the two-pathway model for piRNA 3'-end formation. **b**, **c**, Scatter plots based on piRNAs mapping to germline-dominant TEs ($n = 63$). **b**, Plotted are Z-scores of piRNA 3'/5' coupling in w^{1118} versus $nibbler^{-/-}$ ovaries (orange, Aub-bound antisense piRNA 3' ends to Piwi-bound piRNA 5' ends; blue, Ago3-bound sense piRNA 3' ends to Piwi-bound piRNA 5' ends). **c**, Normalized contributions of Piwi/Aub/Ago3-bound piRNAs to total TE piRNA profiles in w^{1118} versus $nibbler^{-/-}$ ovaries. Stacked bar plots show sum for all analysed TEs.

Fig. 7d). The remaining germline piRNAs in Zucchini and Nibbler co-depleted ovaries populate Aub/Ago3 (Piwi is largely lost; Extended Data Fig. 8a) and exhibit less-defined size profiles, with many piRNAs being abnormally long or short (Extended Data Fig. 8b).

Based on a strong ping-pong signature, Aub/Ago3-mediated slicing defines the 5' ends of Zucchini/Nibbler-independent piRNAs (Extended Data Fig. 8c). A closer look at their mappings provides an

explanation of how their 3' ends are generated (Fig. 4a). In double-depleted ovaries, novel ping-pong pairs emerge between two distantly spaced ping-pong pairs, thereby reducing the cleavage intervals to around 20–30 nt. Also, piRNA 3' ends change from a bell-shaped profile (consistent with Nibbler-mediated exo-resection) to discrete profiles where a single dominating 3' end precedes the 5' end of a flanking piRNA by 1 nt (Fig. 4a). This suggests that two slicer events, spaced by one piRNA length, directly generate 5' and 3' ends of piRNAs. To test this prediction we turned to the piRNA biogenesis reporter with two cleavage sites spaced by 52 nt that is incompatible with piRNA biogenesis in Zucchini and Nibbler co-depleted ovaries (Fig. 4b). Introducing a third central target site re-installs biogenesis for two responder piRNAs, whose 3' ends map precisely to the downstream slicer sites (Fig. 4b).

We systematically analysed Zucchini/Nibbler-independent piRNAs for two characteristic signatures, namely 3'/5' coupling (nucleotide-precision phasing), and 3'/5' ping-pong (presence of complementary piRNA 5' ends 10 nt downstream of piRNA 3' ends). Both signatures—while absent in Zucchini-depleted ovaries—are pronounced in double-depleted ovaries, indicative of tightly spaced ping-pong pairs (Fig. 4c). When piRNAs are grouped into length cohorts, coupling of flanking piRNAs is apparent for all size classes (Fig. 4d). No piRNA coupling is observed in Zucchini-depleted ovaries, as here Nibbler allows ping-pong pairs to be spaced in a larger window (Fig. 4e, Extended Data Fig. 8d).

Zucchini/Nibbler-independent piRNAs bound to Aub or Ago3 retain their respective nucleotide bias of 1U and 10A (Extended Data Fig. 8e). Given the precise piRNA coupling, this explains why slicer/slicer-generated piRNAs display downstream-1U and downstream-10A signatures (Extended Data Fig. 8e). As uridine residues are not spaced in pre-fixed patterns, this requires flexibility on the Argonaute-side to accommodate piRNAs with different lengths. Indeed, while approximately 80% of piRNA species in Zucchini-depleted ovaries are 23–27 nt, this length accounts for only approximately 50% in Zucchini and Nibbler co-depleted ovaries (Fig. 4d, e, Extended Data Fig. 8b). Nucleotide-resolution northern blots confirm the existence of piRNAs as short as 21 nt and as long as 32 nt (Fig. 4f, Extended Data Fig. 9a). While these piRNAs are similarly abundant as the corresponding

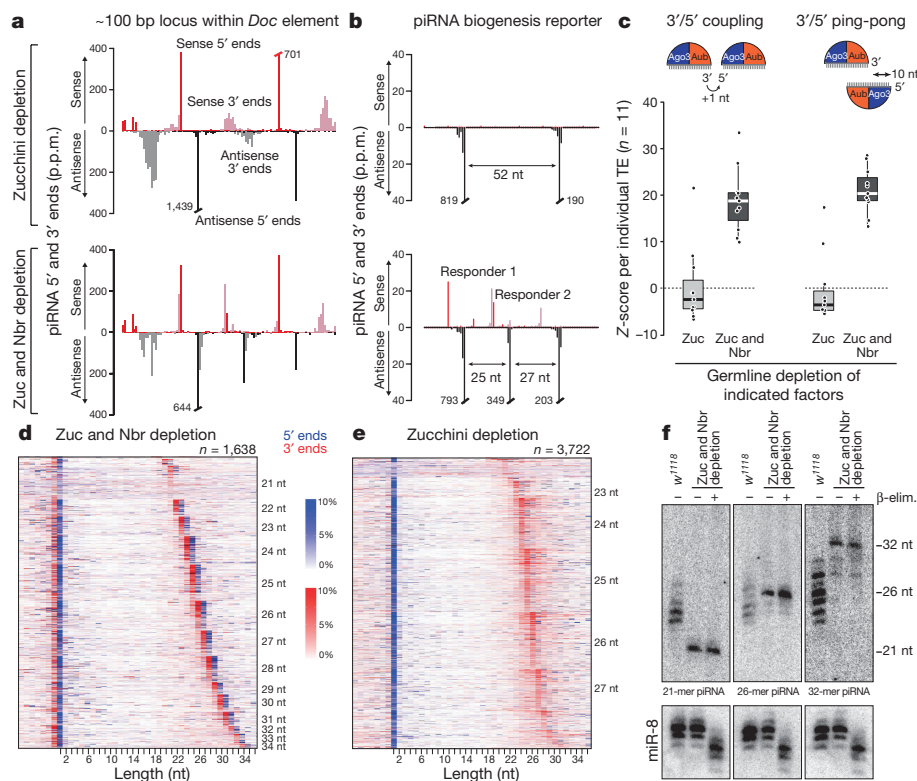


Figure 4 | An Argonaute-only pathway generates piRNAs in the absence of Zucchini and Nibbler. **a**, Normalized 5' and 3' end counts (p.p.m.) of piRNAs isolated from ovaries with indicated genetic background mapping to an approximately 100-nt sequence of *Doc*. **b**, Normalized 5' and 3' end counts (p.p.m.) of piRNAs from Zucchini and Nibbler co-depleted ovaries mapping to a biogenesis reporter with two (top) or three (bottom) piRNA target sites. **c**, Box plots displaying Z-score distributions of 3'/5' coupling and 3'/5' ping-pong for piRNAs from ovaries of indicated genotype ($n = 11$ germline TEs that maintain piRNA production in Zucchini and Nibbler co-depleted ovaries). Midline indicates the median value, box ranges from the first to the third quartile, whiskers are 1.5× the interquartile range. **d**, **e**, Displayed are 5' and 3' ends of piRNAs—isolated from Zucchini and Nibbler co-depleted (**d**) or from Zucchini-depleted (**e**) ovaries—that map within a 40-nt window surrounding ping-pong responder piRNA 5' ends (position 0). Heat maps sorted first by dominant responder piRNA length and then by abundance. The sum of all 3' ends and all 5' ends per line is 100%. **f**, Northern blot analysis of piRNAs with dominant length of 21, 26 or 32 nt in Zucchini and Nibbler co-depleted ovaries.

species in wild-type ovaries, their length is restricted to essentially a single isoform.

When examining slicer/slicer-generated piRNAs for their 2'-O-methylation status, a hallmark of mature piRNA 3' ends^{21,22}, even 32-nt-long piRNAs are methylated (Fig. 4f). This extends to all piRNA sizes (Extended Data Fig. 9b), indicating that 2'-O-methylation can occur independently of Zucchini- or Nibbler-mediated 3'-end formation. As pre-piRNAs with similar length are not methylated in wild-type ovaries (Fig. 1b), Nibbler probably acts faster than Hen1.

Together with recent findings^{8,9,17}, our data provide a blueprint for piRNA 3'-end formation. Two separate exonucleolytic pathways—initiated by endonucleolytic cleavages—are dedicated to pre-piRNA trimming: the Papi/PNLDC1 pathway^{17,18}, and the Nibbler/Mut-7 pathway. This is probably an ancient pathway architecture as—similar to ping-pong²³—all involved nucleases (Zucchini, PNLDC1, and Nibbler/Mut-7) are conserved from sponges to mammals (Extended Data Fig. 10). There are, however, interesting exceptions. (1) Nematodes have lost Zucchini and an entirely different small RNA biogenesis system fuels their PIWI proteins²⁴. In *C. elegans*, PARN trims PIWI-bound small RNAs²⁵, while Nibbler/Mut-7 is required for 22G siRNA biogenesis²⁶. (2) While PARN or PNLDC1 have been sporadically lost in several lineages (for example, fish), only flies have lost both enzymes. We postulate that this central pre-piRNA exonuclease is dispensable in flies as here Zucchini directly forms mature piRNA 3' ends. (3) Only two groups have lost Nibbler: all *Anopheles* species, and several mammals including rodents. Whether this indicates a lesser importance of efficient ping-pong in these species is currently unclear.

The balance between the two 3'-end-generating pathways defines the extent to which precursors are processed into Piwi-bound versus Aub/Ago3-bound piRNAs. Ultimately, this determines the ratio between transcriptional (Piwi) and post-transcriptional silencing (Aub/Ago3). Aub-bound piRNAs, which are antisense-biased, are more abundant than Ago3-bound sense piRNAs. Piwi-bound piRNAs, which are generated in response to Aub/Ago3-initiated triggering, are, however, antisense-biased^{27,28}. Ago3-generated cleavage intermediates must therefore be transferred more efficiently to mitochondria for Zucchini-mediated 3'-end formation than Aub-generated intermediates^{8,9}. Also, Nibbler matures Ago3-bound pre-piRNAs probably more frequently compared to Aub-bound pre-piRNAs, because Aub-bound piRNAs are more abundant than Ago3-bound piRNAs. Consequently, two closely spaced cleavages will be more frequent downstream of Aub than of Ago3.

In the absence of Zucchini and Nibbler, PIWI proteins accommodate 2'-O-methylated piRNAs ranging in length from 20–34 nt. We speculate that the slicer-only pathway potentially represents an ancient small RNA-generating unit, onto which dedicated endo- and exonucleases were added during evolution in order to efficiently generate piRNAs of optimal length.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 23 June; accepted 18 October 2016.

Published online 16 November 2016.

- Ghildiyal, M. & Zamore, P. D. Small silencing RNAs: an expanding universe. *Nat. Rev. Genet.* **10**, 94–108 (2009).
- Malone, C. D. & Hannon, G. J. Small RNAs as guardians of the genome. *Cell* **136**, 656–668 (2009).
- Iwasaki, Y. W., Siomi, M. C. & Siomi, H. PIWI-interacting RNA: its biogenesis and functions. *Annu. Rev. Biochem.* **84**, 405–433 (2015).
- Nishimasu, H. et al. Structure and function of Zucchini endoribonuclease in piRNA biogenesis. *Nature* **491**, 284–287 (2012).
- Ipsaro, J. J., Haase, A. D., Knott, S. R., Joshua-Tor, L. & Hannon, G. J. The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis. *Nature* **491**, 279–283 (2012).

- Brennecke, J. et al. Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* **128**, 1089–1103 (2007).
- Gunawardane, L. S. et al. A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in *Drosophila*. *Science* **315**, 1587–1590 (2007).
- Han, B. W., Wang, W., Li, C., Weng, Z. & Zamore, P. D. Noncoding RNA. piRNA-guided transposon cleavage initiates Zucchini-dependent, phased piRNA production. *Science* **348**, 817–821 (2015).
- Mohn, F., Handler, D. & Brennecke, J. Noncoding RNA. piRNA-guided slicing specifies transcripts for Zucchini-dependent, phased piRNA biogenesis. *Science* **348**, 812–817 (2015).
- Han, B. W., Hung, J. H., Weng, Z., Zamore, P. D. & Ameres, S. L. The 3'-to-5' exonuclease Nibbler shapes the 3' ends of microRNAs bound to *Drosophila* Argonaute1. *Curr. Biol.* **21**, 1878–1887 (2011).
- Liu, N. et al. The exonuclease Nibbler controls 3' end processing of microRNAs in *Drosophila*. *Curr. Biol.* **21**, 1888–1893 (2011).
- Feltzin, V. L. et al. The exonuclease Nibbler regulates age-associated traits and modulates piRNA length in *Drosophila*. *Aging Cell* **14**, 443–452 (2015).
- Wang, H. et al. Antagonistic roles of Nibbler and Hen1 in modulating piRNA 3' ends in *Drosophila*. *Development* **143**, 530–539 (2016).
- Vourekas, A. et al. Mili and Miwi target RNA repertoire reveals piRNA biogenesis and function of Miwi in spermiogenesis. *Nat. Struct. Mol. Biol.* **19**, 773–781 (2012).
- Kawaoka, S., Izumi, N., Katsuma, S. & Tomari, Y. 3' end formation of PIWI-interacting RNAs in vitro. *Mol. Cell* **43**, 1015–1022 (2011).
- Olivieri, D., Senti, K. A., Subramanian, S., Sachidanandam, R. & Brennecke, J. The cochaperone shutdown defines a group of biogenesis factors essential for all piRNA populations in *Drosophila*. *Mol. Cell* **47**, 954–969 (2012).
- Izumi, N. et al. Identification and functional analysis of the Pre-piRNA 3' trimmer in silkworms. *Cell* **164**, 962–973 (2016).
- Saxe, J. P., Chen, M., Zhao, H. & Lin, H. Tdrkh is essential for spermatogenesis and participates in primary piRNA biogenesis in the germline. *EMBO J.* **32**, 1869–1885 (2013).
- Sato, K. et al. Krimper enforces an antisense bias on piRNA pools by binding AGO3 in the *Drosophila* germline. *Mol. Cell* **59**, 553–563 (2015).
- Webster, A. et al. Aub and Ago3 are recruited to Nuage through two mechanisms to form a ping-pong complex assembled by Krimper. *Mol. Cell* **59**, 564–575 (2015).
- Horwich, M. D. et al. The *Drosophila* RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. *Curr. Biol.* **17**, 1265–1272 (2007).
- Saito, K. et al. Pimet, the *Drosophila* homolog of HEN1, mediates 2'-O-methylation of Piwi-interacting RNAs at their 3' ends. *Genes Dev.* **21**, 1603–1608 (2007).
- Grimson, A. et al. Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. *Nature* **455**, 1193–1197 (2008).
- Sarkies, P. et al. Ancient and novel small RNA pathways compensate for the loss of piRNAs in multiple independent nematode lineages. *PLoS Biol.* **13**, e1002061 (2015).
- Tang, W., Tu, S., Lee, H. C., Weng, Z. & Mello, C. C. The RNase PARN-1 trims piRNA 3' ends to promote transcriptome surveillance in *C. elegans*. *Cell* **164**, 974–984 (2016).
- Gu, W. et al. Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the *C. elegans* germline. *Mol. Cell* **36**, 231–244 (2009).
- Senti, K. A., Jurczak, D., Sachidanandam, R. & Brennecke, J. piRNA-guided slicing of transposon transcripts enforces their transcriptional silencing via specifying the nuclear piRNA repertoire. *Genes Dev.* **29**, 1747–1762 (2015).
- Wang, W. et al. Slicing and binding by Ago3 or Aub trigger Piwi-bound piRNA production by distinct mechanisms. *Mol. Cell* **59**, 819–830 (2015).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank all laboratory members for help and discussions, P. Duchek, J. Gokceade and K. Meixner for generating fly lines, M. Novatchkova for help on the conservation analysis of nucleases, the VBCF NGS facility for sequencing, and the MFPL monoclonal facility for the Nibbler antibody. This work was supported by the Austrian Academy of Sciences, the European Community's 7th Framework Program (ERC-StG-260711; ERC-StG-338252), the Austrian Science Fund (Y510-B12; F4303-B09; W12-7-B09; Y733-B22), and an HFSP postdoctoral fellowship to F.M.

Author Contributions F.M. made the initial observation that Nibbler trims Ago3-bound piRNAs, J.S. and R.H. did all experiments and did the computational analysis with the help of D.H. All authors designed the experiments and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.B. (julius.brennecke@imba.oeaw.ac.at) and S.A. (stefan.ameres@imba.oeaw.ac.at).

Reviewer Information *Nature* thanks B. Czech and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Fly husbandry and strains. Flies were kept at 25 °C. Short hairpin RNA (shRNA) fly lines were crossed to the maternal triple driver (MTD)-GAL4 line (#31777; Bloomington stock centre) to drive expression of the shRNA in the germ line. shRNA constructs for double depletion of Zucchini-Nibbler and Zucchini-Papi were inserted into attP2²⁹. TriP.GL00111 line was used (#35227; Bloomington) for depleting Zucchini alone. GFP reporter constructs and GFP-tagged BAC rescue constructs were inserted into the attP40 landing site. A *nibbler* and a *papi* allele with frame-shift mutations were generated by CRISPR/Cas9 using guide RNAs TGACGCCACCTTGGACGCAA and CGAGCCGCCTTAACCGCATC, respectively, as previously described³⁰. N-terminally Flag-tagged *nibbler* allele was generated using a guide RNA CACGGGAAACCCGTGAGAAA. The resulting allele has an insertion of IDYKDHGDYKDHIDYKDD after the start codon. *w¹¹¹⁸* strain was used as a wild-type control throughout the study except the analysis of TE expression from ovaries using RNA sequencing where an shRNA line against *white* crossed to MTD-GAL4 was used as a control²⁷. Flies were aged for 6 days and kept on apple juice agar plates supplemented with yeast paste to ensure consistent ovarian morphology.

Construction of shRNA expression vectors. To achieve simultaneous knockdown of two genes with a single shRNA construct, we modified the Valium20 vector³¹ and inserted two tandem shRNA sequences. The two hairpins are separated by the sequence that spans the miR-6-3 and miR-6-2 hairpins in the genome in order to maximize efficient processing (referred to as miR-6 backbone in Extended Data Fig. 2a). *NheI*/*EcoRI* and *AgeI*/*SphI* sites were used to clone the two shRNAs oligos. The modified shRNA expression cassette (restriction sites used for cloning of shRNA oligos are underlined and the miR-6 spacer sequence is in bold):

TCAGCCGCTAGCATGGACGTTTCATGGATTCTAAACGGATTACCAG
GGATTTTCAGTCGATGTGAATTCAGGCGAGCACGGCCAATTCCAACGA
TTTGTCATTTGTGGCAGCATTTGTGTCACTCAGTGGGAAAATTGAA
AATTTGATTCAGCCACCGGTTGCATAAGGCTATGAAGAGATACGCCCT
GCATGCAGGCGAG

The shRNA oligo sequences are:

Zucchini shRNA forward (guide miR sequence is underlined):

CTAGCATGCAGCACTTGATGCACAAACAATAGTTATATTCAAGCATAT
TGTTGTGCATCAAGTTCGTGGCG

Zucchini shRNA reverse:

AATTCGCCACGAACCTTGATGCACAAACAATATGCTTGAATATAACTATT
GTTGTGCATCAAGTTCGTGACTG

Nibbler shRNA forward (guide miR sequence is underlined):

CCGGTAGTATGGTCACTGATCTCAGTGTATAGTTATATTCAAGCATATA
CACTGAGATCACTGACCATGCGCATG

Nibbler shRNA reverse:

CGCATGGTCAGTGTCTCAGTGTATATGCTTGAATATAACTATACACTG
AGATCACTGACCATATA

Papi shRNA forward (guide miR sequence is underlined):

CCGGTAGTTCGACATATCCTAGATCCTAATAGTTATATTCAAGCATATT
AGGATCTAGGATATGTGAGCGCATG

Papi shRNA reverse:

CGCTCGACATATCCTAGATCCTAATATGCTTGAATATAACTATTAGGAT
CTAGGATATGTGCAACTA

GFP-tagged Pacman rescue constructs. Recombineering of Pacman³² rescue constructs was as described³³. *Papi* (Pacman clone CH322-41G09) and *Zucchini* (CH322-41M17) were tagged with GFP-Flag tags C-terminally, and *Nibbler* (CH322-18I04) was tagged N-terminally.

Generation of piRNA reporter constructs. All piRNA reporter constructs were generated as described previously⁹ using the following sequences (target sites for endogenous piRNAs are marked in bold):

Reporter with two target sites (Fig. 1d):

GGACGAGCTGTACAAGTAATGATCGAGTCAGATAAGCCGTATCAATC
AAAGGTTATCAAAGCACTGATGATACAAGAAATCGCCTCGGTTT
AATGAAATAACTCCAAGAATGCTCATTGAAATGATACTGTCAAATGTT
TCACAGGAAACTACTCAAATATAACCAAAATGAAGCAGTCATCAAGT
GAACGAAAGAAATGAAACATGACACGTGATCAACTCGAGCGCATAGA
AATTAATTTATAAATTGCAAAACCTAATTACGTAGCTAAACAAAAACC
CAAAAGAAATAATGTAAGCAAAAAGAAATTAATCGTCATAATCGTCAC
CAAACCTTTTGTATTGGCGAGCATTTTCATAGATGTTAAATTTTCTAATT
CGAGAATCCCAGGAGGATAAGCGATAGGGATGATCAGAA

Reporter with two target sites (Extended Data Fig. 2f):

GGACGAGCTGTACAAGTAATGATCGAGTCAGATAAGCCGTATCAAT
TTCAAAGGTTATCAAAGAGCAACTGAAAACCCGTATAGTGACTATA

CAGCAACTCCCACCTTAAAAAGGAAATATTCATGAACCTGCTTTAGCC
TAGGAAACTACTCAAATATAACCAAAATGAAGCAGTCATCAAGTGAA
CGAAAGAAATGAAAACATGACACGTGATCAACTCGAGCGCATAGAAAT
TAATTTATAAATTGCAAAACCTAATTACGTAGCTAAACAAAAACCCAAA
AGAATAATGTAAGCAAAAAGAAATTTAATCGTCATAATCGTCACCAAA
CCTTTTGTATTGGCGAGCATTTTCATAGATGTTAAATTTTCTAATTTCGAGA
ATCCCAGGAGGATA AGCGATAGGGATGATCAGAA

Reporter with three target sites (Fig. 4b):

GGACGAGCTGTACAAGTAATGATCGAGTCAGATAAGCCGTATCAAT
TTCAAAGGTTATCAAAGAGCAACTGAAAACCCGTATAGTGACTAAA
TACCTACTCCCACCTTAAAAAGGAAATATTCATGAACCTGCTTTAGCC
TAGGAAACTACTCAAATATAACCAAAATGAAGCAGTCATCAAGTGAA
CGAAAGAAATGAAAACATGACACGTGATCAACTCGAGCGCATAGAAAT
TAATTTATAAATTGCAAAACCTAATTACGTAGCTAAACAAAAACCCAAA
AGAATAATGTAAGCAAAAAGAAATTTAATCGTCATAATCGTCACCAAA
CCTTTTGTATTGGCGAGCATTTTCATAGATGTTAAATTTTCTAATTTCGAG
AATCCCAGGAGGATAAGCGATAGGGATGATCAGAA

Generation of Nibbler antibody. Purified His-tagged full-length Nibbler protein was used to generate the mouse anti-Nibbler monoclonal antibody.

Antibodies used. The following primary antibodies were used. Mouse monoclonal anti-Piwi (8C-E4)²⁷ (western: 1/1,000; used for IP), rabbit anti-Piwi⁶ (IF: 1/500), mouse monoclonal anti-Ago3 (7B4-C2)²⁷ (western: 1/1,000; used for IP), mouse monoclonal anti-Ago3 (5H12-G12)²⁷ (IF: 1/30), mouse monoclonal anti-Nibbler (IF: 1/400; western: 1/2,000), mouse monoclonal anti-Aub (8A3-D7)²⁷ (IF: 1/500; western: 1/1,000; used for IP), mouse anti-ATP-synthase 5A (abcam 14748; IF: 1/2000; western: 1/20,000).

Secondary antibodies used were: goat-anti-mouse HRP (Invitrogen), Alexa Fluor 488 anti-mouse (Thermo Fisher), Alexa Fluor 568 anti-mouse (Thermo Fisher), Alexa Fluor 647 anti-mouse (Thermo Fisher), Alexa Fluor 568 anti-rabbit (Thermo Fisher).

Immunofluorescence staining. Ovaries were dissected into 300 µl PBS containing 4% PFA and fixed for 25 min at room temperature (RT). Tissue was permeabilized 3 times with PBX (1 × PBS, 0.3% TritonX-100) and blocked with BBX (1 × PBS, 0.3% TritonX-100, 0.1% BSA) for 30 min. 200 µl of primary antibodies diluted in BBX were added and ovaries were incubated while rotating over night at 4 °C. After washing 3 times with PBX, the ovaries were incubated with the respective secondary antibodies diluted in BBX rotating overnight at 4 °C. Subsequently, ovaries were washed three times with PBX before mounting. To the second washing step DAPI was added. Images were taken on a Zeiss LSM-780 Axio Imager confocal-microscope using a 40 ×/1.3 EC plan-neofluar Oil DIC objective. All images were processed using ImageJ.

Immunoprecipitation of PIWI-family proteins for piRNA sequencing. For each genotype, 600 µl of ovaries were dissected into 1 × PBS on ice. 1.5 ml of 1 × RIPA buffer were added (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% TritonX-100, 0.1% SDS, 0.1% Na-deoxycholate, 1 mM EDTA, 0.1 mM Pefabloc (Roche)) and the tissue was homogenized, using a glass tissue homogenizer, on ice. The lysate was cleared by centrifugation and diluted with 3 ml of IP dilution buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl). Antibodies were coupled to M280 sheep anti-mouse IgG Dynabeads (Life Technologies). For Piwi and Aub IP, 150 µl of beads coupled to the respective antibody were mixed with 1.5 ml of lysate. For Ago3 IP, 300 µl of beads coupled to Ago3 antibody were mixed with 3 ml of lysate. Antibodies used are indicated above. Lysates were incubated rotating at 4 °C overnight. Subsequently, the beads were captured and washed seven times with IP wash buffer (50 mM Tris-HCl pH 7.5, 500 mM NaCl, 2 mM MgCl₂, 10% glycerol, 1% Empigen). For Piwi IP, 150 mM NaCl was used instead of 500 mM NaCl. The bound RNA was extracted using acid-phenol:chloroform (Ambion) followed by ethanol precipitation. To visualize the extracted RNA during the cloning process 10% of it was labelled with [γ -³²P]-ATP.

Co-immunoprecipitation of Flag-Nibbler. 100 µl each of freshly dissected ovaries from 1–2-day-old *w¹¹¹⁸* and *flag-nibbler* homozygous females were homogenized in lysis buffer (30 mM HEPES-KOH pH 7.4, 500 mM NaCl, 2 mM MgCl₂, 0.5% v/v Triton-X, 0.2 mM DTT, and 1 × cComplete, EDTA-free protease inhibitor cocktail (Roche)) using a glass tissue homogenizer, on ice. The lysates were cleared by centrifugation and incubated with 30 µl of anti-Flag M2 agarose beads (Sigma) for 4 h at 4 °C. Beads were captured and washed 5 times with the lysis buffer. IP fraction was eluted by incubating the beads with 30 µl of lysis buffer containing 0.5 µg µl⁻¹ 3 × Flag peptides for 15 min at RT with gentle shaking.

2S rRNA depletion from total RNA. For the depletion of 2S rRNA from 10 µg of total ovarian RNA, 100 µl slurry of Myone Streptavidin C1 beads (Invitrogen) were used. The beads were washed twice with 500 µl 0.5 × SSC. After washing, the beads were re-suspended in 300 µl 0.5 × SSC and 1 µl of 100 µM 2 s-rRNA-as-oligo (Bio-AGTCTTACAACCTCAACCATATGTAGTCCAAGCAGCACT) were added and mixed. This mixture was incubated on ice for 30 min, followed by

a wash with 500 μ l of 0.5 \times SSC to remove unbound 2-s-rRNA-as-oligo. After resuspension of the beads in 100 μ l, they were incubated at 65 °C for 5 min. The RNA was denatured for 5 min at 80 °C for 5 min and mixed with the beads, followed by incubation at 50 °C for 1 h. Fraction not bound to the beads was collected and the RNA was ethanol precipitated and served as input for the small RNA cloning procedure.

Small RNA cloning and sequencing. Small RNA libraries were prepared as described previously³⁴. In brief, total RNA from respective ovaries was isolated using TRIzol, and was subjected to 2S-depletion³⁵. Oxidation was done by incubating 2S-depleted total RNA in borate buffer (pH 8.6) containing 25 mM NaIO₄ at room temperature for 30 min. RNA samples from total RNA and from IP experiments were resolved on a denaturing polyacrylamide gel and small RNAs ranging from 18–40 nt were excised and recovered. These were subsequently ligated to 3' and 5' adapters containing four random nucleotides at both ends to reduce ligation biases. Ligated RNA was reverse-transcribed, PCR amplified and the libraries were sequenced on an Illumina HiSeq 2500 machine in single-read 50 mode. IP libraries of the Zucchini/Nibbler-depleted flies were sequenced in single-read 100 mode.

RNA sequencing. PolyA+ RNA-seq was performed as described in Mohn *et al.*³⁶ using NEBNext Ultra directional RNA library prep kit for Illumina (NEB) and libraries were sequenced on Illumina HiSeq 2500 in paired-end 125 mode³⁶. Six bases at each end of the reads were trimmed and the remaining part was split into three reads (37, 38, 38 nt). Reads were mapped against TE consensus sequences using Bowtie (0.12.9) allowing up to three mismatches²⁷. RPKM for each TE was calculated as reads per million genome mapped reads per kb. TEs that were expressed more than 1 RPKM in at least one of the libraries in the comparison (group1: *w¹¹¹⁸* and *nibbler^{-/-}*, group2: control depletion, zucchini depletion and zucchini and nibbler depletion) were included for the analysis.

Northern blot. Small RNAs were enriched from 50 μ g of total RNA using size selection on a 12% denaturing urea-polyacrylamide gel. Oxidation was performed as mentioned above. β -elimination was done after the oxidation by adding 50 mM f.c. NaOH and incubating for 90 min at 45 °C. The protocol for northern blotting was modified from Pall *et al.*³⁷.

In brief, a 15% denaturing urea-polyacrylamide gel (0.4 mm thick) was used to run the size-selected RNA samples. The gel was blotted to a Hybond-NX membrane (Amersham) in a semi-dry blotting apparatus. This was followed by chemical cross-linking using 0.373 g 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) (Sigma) powder in 12 ml 1 \times methylimidazole (Sigma) at 60 °C for 1 h. Afterwards the membrane was pre-hybridized in 10–15 ml (depending on membrane size) church buffer (1 mM EDTA, 0.5 M Na₂HPO₄/NaH₂PO₄ pH 7.2, 7% SDS) at 65 °C for at least 20 min. DNA oligonucleotide probes were labelled using [γ -³²P]-ATP (PerkinElmer). The radioactively-labelled probe was added to the church buffer and incubated overnight at 32 °C. The membrane was subsequently washed 3 times with 15 ml of 1 \times SSC + 0.1% SDS. Finally, a phosphor-storage screen was exposed to the membrane and developed using a phosphor imager. All images were processed using ImageJ.

The following DNA oligos were used as probes:

miR-34 (Extended Data Fig. 3d):
AACCAGCTAACCACACTGCCA
miR-8 (Figs 1b, 2a, b and 4f):
GACATCTTACCTGACAGTATTA
piRNA enriched in Ago3 (Fig. 2a):
GCGATTTCTTGGGTTTCAGTTGCT
piRNA enriched in Aubergine (Fig. 2b):
TCTTGGAGCTAAGTCTTTCGTA
21-nt-long piRNA (Fig. 4f):
TCTAGATTGGCTGCTATTA
26-nt-long piRNA (Fig. 4f):
AAGCTACTGAAGTCATACCTATA
32-nt-long piRNA (Fig. 4f):
AGAACAACTGGCCAAGGTATCAA
piRNA precursor (Fig. 1b):
AGTCTGGAGTTCAAAGCTCTTCTA

Processing of small RNA libraries. Processing and mapping of sequenced small RNA reads was carried out as described in Mohn *et al.*⁹. In brief, raw reads were trimmed off the 3' linker sequence (AGATCGGAAGAGCACACGTCT) and the four random nucleotides at each end were removed. Processed reads were mapped to the *Drosophila* genome (dm3, release 5.55) using Bowtie (0.12.9) allowing zero mismatches. Reads that mapped to genomic regions annotated as TE were used for the subsequent analyses (filtered reads). Libraries from total RNA were normalized to 1 million miRNA reads. Normalization of PIWI-IP libraries was as described previously⁹. R (version 2.15.3) was used for statistical and graphical analyses unless otherwise indicated.

TEs that were included in the analyses. 63 TEs that fulfilled the following criteria were defined as germline-enriched TEs. 1) TEs that produced more than 1,000 p.p.m. of piRNAs (sum of sense and antisense piRNAs) in wild-type ovaries; 2) germline-derived Piwi-bound piRNAs constitute more than 85% of total Piwi-bound piRNAs²⁷. 63 TEs were grouped into two groups based on the ratio of summed TE mapping reads of normalized Piwi/Aub/Ago3 IP libraries in Zucchini-depleted ovaries relative to the control depletion (group1: >0.2, *n* = 19, group2: <0.2, *n* = 44, used in Extended Data Figs 1b and 6d). 11 out of 63 TEs that fulfilled the following criteria were used for the coupling/ping-pong analyses in Fig. 4c and Extended Data Fig. 8c. 1) TEs that produced more than 10% of piRNAs in Zucchini and Nibbler-depleted ovaries compared to the control depletion; 2) TEs with a canonical ping-pong Z-score (see below) greater than 15.

Statistical analysis of size profiles. 1,000 TE-mapping reads were randomly sampled 100 times. The size distribution of these 1,000 reads was compared between genotypes using a *t*-test. The mean *P* value of 100 comparisons was calculated to determine the significance.

Linkage analysis. Linkage calculation was adapted from Webster *et al.*²⁰. First, filtered small RNA reads were mapped to TE consensus sequences²⁷ allowing three mismatches. Second, 5' and 3' ends of mapped reads were counted at each nucleotide position on both strands. For each linkage analysis, the statistical significance of observing a certain linkage was calculated as follows.

- (1) For example, for the calculation of canonical ping-pong, piRNA 5' ends were randomly subsampled for a fixed number of times from antisense and sense mapped piRNAs. We sampled for the number of a fifth of TE length (bp). For instance, if the size of a TE is 5,000 bp, we sampled 5' ends 1,000 times. A fixed number of samplings allows for a fair comparison between libraries of different sequencing depths. The shuf function in bash 4.2.37(1) was used to obtain random numbers.
- (2) We counted the number of 5' ends of antisense piRNAs that find at least one 5' end of sense piRNA at 10-nt off-set (*N*).
- (3) We computationally repeated (1) and (2) for 500 times and calculated the mean of N_{mean} .
- (4) Suppose two sets of random 5' ends, the probability of an antisense 5' end to find a ping-pong partner *p* is $1 - (1 - 1/L)^L$ where *L* is the length of TE. The expected number of *N* from a random set (N_{expected}) is $L/5 \times p$.
- (5) Z-score was calculated as $(N_{\text{mean}} - N_{\text{expected}})/\sigma$ where the standard deviation σ is $(L/5 \times p \times (1 - p))^{0.5}$.

We used the following sets of sequencing reads for calculation. Canonical ping-pong: 5' ends of antisense and sense mapped reads from total libraries (10-nt off-set, Extended Data Figs 1b and 8c). 3'/5' coupling: 3' ends and 5' ends of antisense mapped reads from total libraries (1-nt off-set, Extended Data Figs 1b and Fig. 4c). 3'/5' ping-pong: 3' ends of antisense mapped reads and 5' ends of sense mapped reads (10-nt off-set, Extended Data Figs 1b and Fig. 4c).

Mapping small RNA reads to reporter constructs. Responder and trigger piRNA reads were determined by mapping the total small RNA reads including genome-unmapped reads to reporter construct sequences using bowtie allowing zero mismatches. To determine trigger piRNA levels, only the first 20 nt of sequencing reads were used in order to account for 3' heterogeneity. Mapped reads were normalized to 1 million miRNA reads.

Measure the definition of 3' ends. The analysis was adapted from Mohn *et al.*⁹. In brief, piRNA 5' end positions where the downstream 20–35 nt window had equal number of mappings in the reference genome were selected as unambiguous 5' end positions. We only included the reads whose 5' ends are defined as unambiguous (analysable reads). These analysable reads were collapsed on 5' ends yielding the counts of each length per given 5' end. 5' ends that have more than 20 raw counts as well as more than 2 counts normalized to one million TE-mapping reads (p.p.m.) were included in the analysis. These cut-offs allowed the comparison of different sequencing depths as well as the confident assessment of 3'-end variants. Finally, the fraction of the counts representing the dominant length per 5' end was calculated as a percentage (definition). For the analysis of somatic piRNAs, piRNAs that mapped to soma-enriched 1 kb tiles were used with a cut-off of soma-index greater than 8³⁶.

Nucleotide analysis. Only piRNA reads that passed the requirements for 3'-end analysis (see above) were used for the analysis. The entire analysis is sequence-based, not read-based (cloning frequency is therefore not considered). Nucleotide windows surrounding the respective positions were extracted using get-fasta from the fastX-toolkit (Hannon laboratory). For each 5' end, the dominant 3' end position was used to extract the surrounding sequences. When indicated, the 5' ends were binned into ten groups containing an equal number of 5' ends according to the definition of their 3' ends. The nucleotide signatures were generated using weblogo 3.4 and Prism 6 was used for visualization.

Contribution analysis for Nibbler and Zucchini. The relative contribution of Nibbler and Zucchini in forming the 3' ends of Aub- or Ago3-bound piRNAs was modelled as follows:

Analysable piRNA 5' ends (see above) with at least 20 p.p.m. in the respective IP libraries of all three genotypes (w^{1118} , *nibbler*^{-/-}, and Zucchini depletion) were included in the analysis. First, the 3'-end composition of individual 5' ends in a window of 20–30 nt of length was determined. The 3'-end profile of piRNAs from *nibbler*^{-/-} and Zucchini-depletion libraries were summed in different ratios (0–100% and the reciprocal values) to generate a combined profile. Second, the combined profile was compared to the w^{1118} profile using a linear regression model. The R^2 value was used to assess the best fit. Finally, the ratios of Nibbler versus Zucchini contribution were extracted at the maximal R^2 value.

Heat maps for 3'-to-5' coupling. Heat maps were generated as previously described⁹. In brief, filtered reads (see above) from total small RNA libraries were mapped to TE consensus sequences allowing up to three mismatches. We selected pairs of complementary reads with 10-nt 5' overlap (ping-pong piRNA pairs). Ping-pong piRNAs that were cloned greater than 10 p.p.m. and more than 2% of the most abundant piRNA per individual TEs were selected for the analysis. We extracted the counts of 5' and 3' ends mapped in the window of ± 100 nt for each ping-pong piRNA (two instances per pair) using annotatePeaks (Homer v.4.5) and transformed the counts at each position to a percentage value (read count table). The read count tables per ping-pong piRNA were sorted by the length of the dominant ping-pong responder piRNA and by the relative abundance of the ping-pong responder piRNA in the window, and were assembled into heat maps. The distribution of 5'-to-5' distance of flanking ping-pong piRNAs was calculated by summing the read count table of all ping-pong piRNAs in the respective genotypes (Extended Data Fig. 8d).

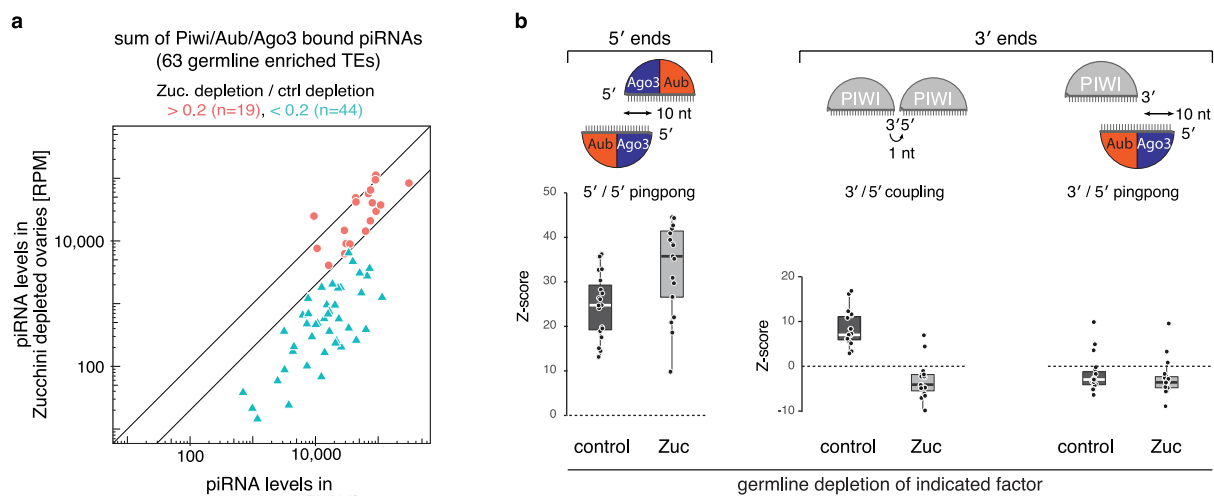
Orthologue search. Orthologues for Zucchini/MitoPLD, PARN, PNLDC1, and Nibbler/Mut7 were first searched using orthoDB v.9 against metazoan species (<http://www.orthodb.org/>)³⁸. Orthologue groups (EOG091G0FI5: Zucchini/MitoPLD, EOG091G04CM: PARN/PNLDC1, EOG091G04KJ: Nibbler/Mut7) were analysed. Protein entries that were not annotated with the following pfam domains were excluded: Zucchini/MitoPLD: PLDC_2 (PF13091), PARN/PNLDC1: CAF1 (PF04857), Nibbler/Mut7: DNA_pol_A_exo1 (PF01612). Pfamscan (<http://www.ebi.ac.uk/Tools/pfa/pfamscan/>) was used to search for the domains (pfam-A, E-value: 0.1). We defined PARN as proteins that have RNA_bind (PF08675) and/or RRM_7 (PF16367) and PNLDC1 as proteins that do not have these annotations. Species that did not have orthologues in orthoDB were further searched for orthologues first by BLASTp (NCBI) against non-redundant protein sequences and second by tBLASTn (NCBI) against whole genome shotgun sequences using the default cut-offs. Hits from BLASTp and tBLASTn were reciprocally surveyed

for predicting the query proteins as the most confident homologue. All hits were combined and displayed in a phylogenetic tree (either presence or absence) using iTOL (<http://itol.embl.de/>)³⁹.

Code availability. The code that underlies the computational analyses in this manuscript is available at https://gitlab.com/groups/Genetic_and_mechanistic_diversity_of_piRNA_3prime_end_formation.

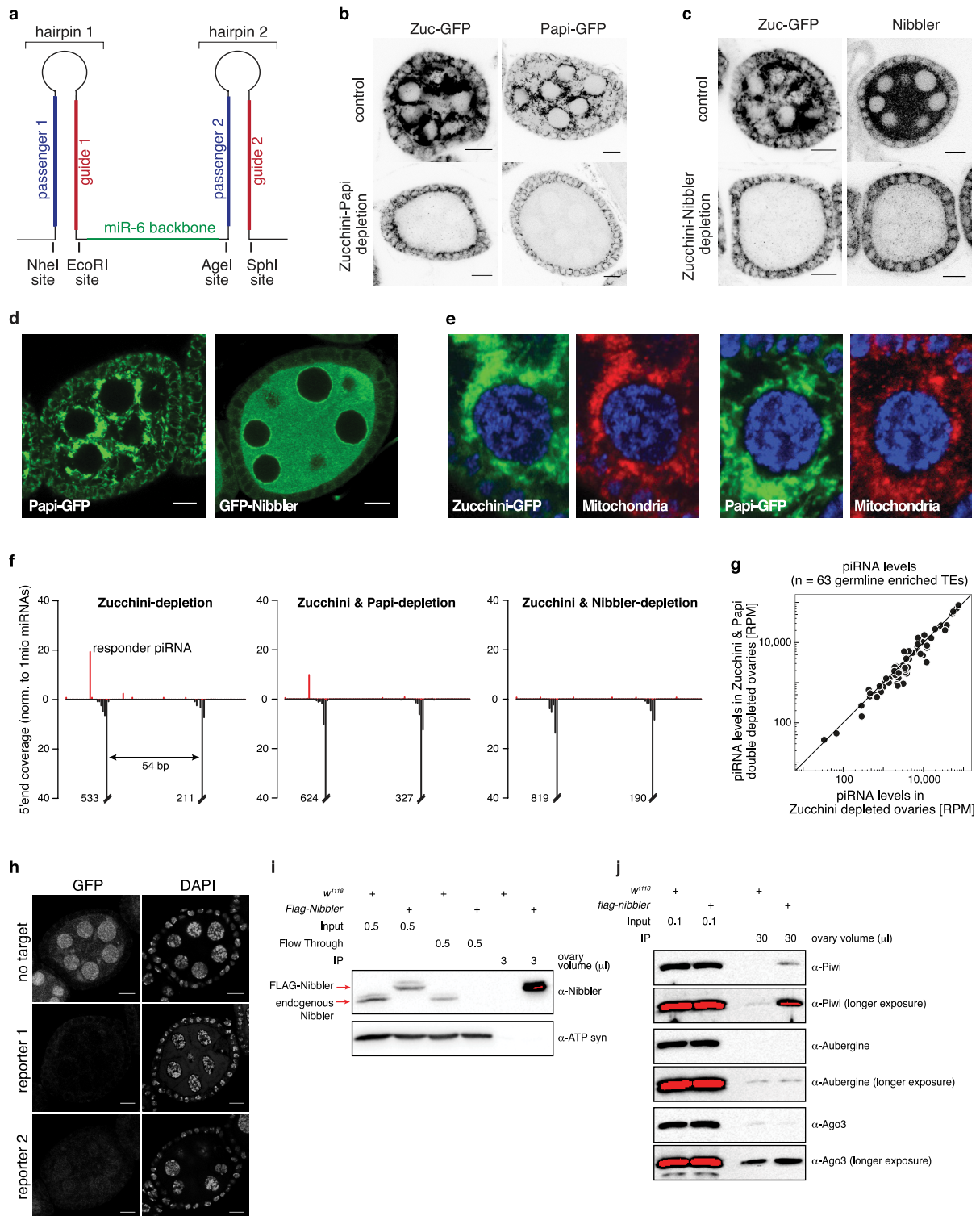
Data availability statement. All sequence data that support the findings of this study (see SI) have been deposited in NCBI GEO with the accession code GSE83698. All fly lines used in this study are available from VDRC (<http://stockcenter.vdrc.at/control/main>). Source data for all gel images are provided with the paper (see Supplementary Information).

29. Markstein, M., Pitsouli, C., Villalta, C., Celniker, S. E. & Perrimon, N. Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. *Nat. Genet.* **40**, 476–483 (2008).
30. Gokcezade, J., Sienski, G. & Duchek, P. Efficient CRISPR/Cas9 plasmids for rapid and versatile genome editing in *Drosophila*. *G3 (Bethesda)* **4**, 2279–2282 (2014).
31. Ni, J. Q. *et al.* A genome-scale shRNA resource for transgenic RNAi in *Drosophila*. *Nat. Methods* **8**, 405–407 (2011).
32. Venken, K. J. *et al.* Versatile P[acman] BAC libraries for transgenesis studies in *Drosophila melanogaster*. *Nat. Methods* **6**, 431–434 (2009).
33. Ejsmont, R. K. *et al.* Recombination-mediated genetic engineering of large genomic DNA transgenes. *Methods Mol. Biol.* **772**, 445–458 (2011).
34. Jayaprakash, A. D., Jabado, O., Brown, B. D. & Sachidanandam, R. Identification and remediation of biases in the activity of RNA ligases in small-RNA deep sequencing. *Nucleic Acids Res.* **39**, e141 (2011).
35. Seitz, H., Ghildiyal, M. & Zamore, P. D. Argonaute loading improves the 5' precision of both MicroRNAs and their miRNA* strand in flies. *Curr. Biol.* **18**, 147–151 (2008).
36. Mohn, F., Sienski, G., Handler, D. & Brennecke, J. The rhino-deadlock-cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in *Drosophila*. *Cell* **157**, 1364–1379 (2014).
37. Pall, G. S., Codony-Servat, C., Byrne, J., Ritchie, L. & Hamilton, A. Carbodiimide-mediated cross-linking of RNA to nylon membranes improves the detection of siRNA, miRNA and piRNA by northern blot. *Nucleic Acids Res.* **35**, e60 (2007).
38. Kriventseva, E. V. *et al.* OrthoDB v8: update of the hierarchical catalog of orthologs and the underlying free software. *Nucleic Acids Res.* **43**, D250–D256 (2015).
39. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* **23**, 127–128 (2007).



Extended Data Figure 1 | 3' ends of Zucchini-independent ping-pong piRNAs are formed by an exonuclease. a, Scatter plot showing the fold change in piRNA levels (for 63 germline-dominant TEs) in Zucchini-depleted compared to control ovaries (calculated as sum of normalized Piwi/Aub/Ago3-bound piRNAs). TEs were grouped into robust (red) and sensitive (blue) on the basis of piRNA loss (threshold = $5\times$ loss). **b,** Box

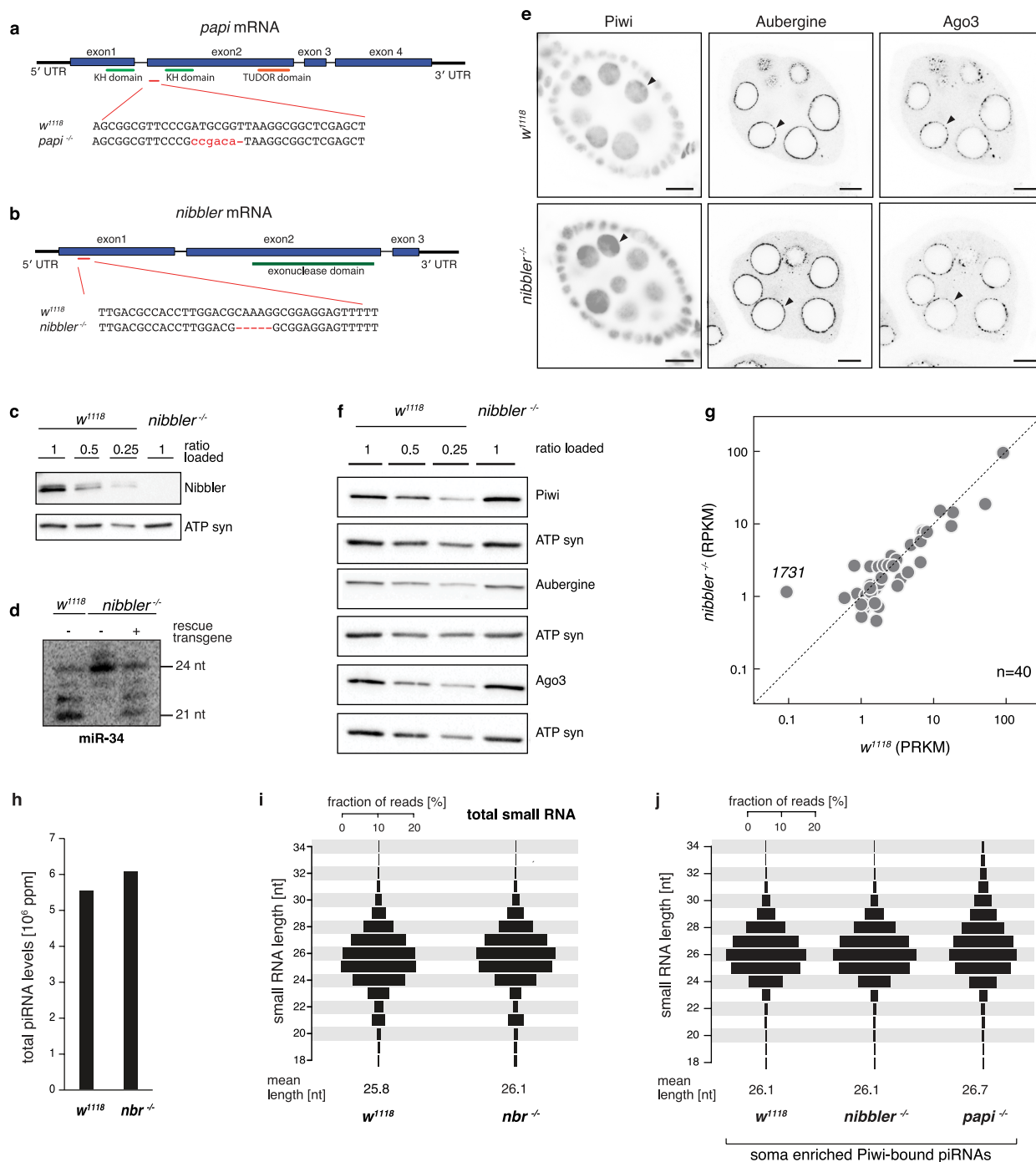
plots displaying the Z-scores of canonical 5'/5' ping-pong, 3'/5' coupling, and 3'/5' ping-pong for piRNAs isolated from ovaries of indicated genotype (for the 19 robust germline-enriched TEs that maintain piRNA production in Zucchini-depleted ovaries; defined in panel a). Midline indicates the median value, box ranges from the first to the third quartile, whiskers are $1.5\times$ the interquartile range.



Extended Data Figure 2 | See next page for caption.

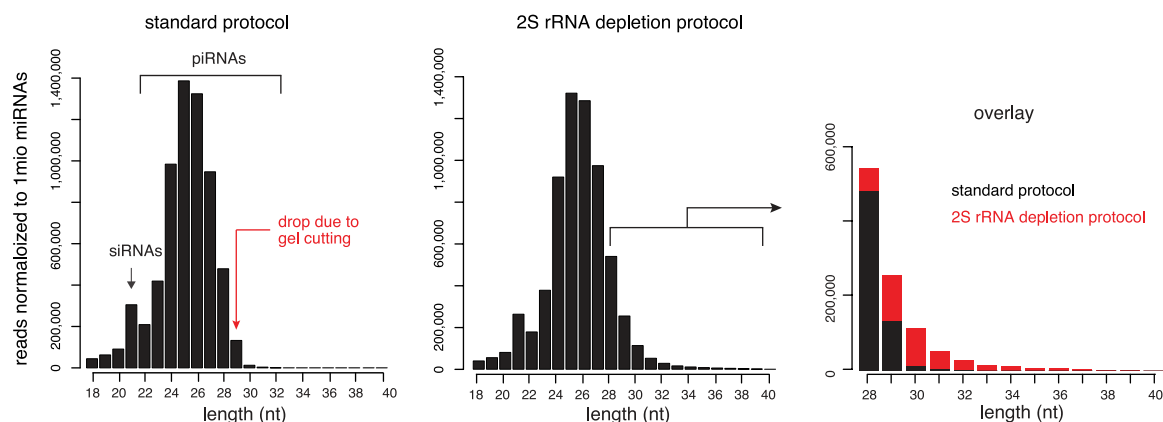
Extended Data Figure 2 | Nibbler, but not Papi, is required for the generation of piRNAs in the absence of Zucchini. **a**, Design of the short hairpin expression cassette that allows the simultaneous RNAi-mediated knockdown of two genes in a tissue specific manner. shRNAs are separated by the miR-6 backbone and can be cloned via indicated restriction sites. **b, c**, Confocal sections of egg chambers (scale bars, 10 μ m) of indicated genotype expressing GFP-tagged Zucchini (**b, c**) and GFP-tagged Papi (**b**), showing the efficient shRNA-mediated knockdown of Zucchini and Papi (**b**), or Zucchini and Nibbler (**c**) in the germ line. Nibbler was detected using a monoclonal antibody. **d**, Confocal images showing the localization of GFP-tagged Papi and Nibbler in *Drosophila* egg-chambers (scale bars, 10 μ m). Functionality of GFP-Nibbler is demonstrated in Extended Data Fig. 3d. **e**, Confocal sections through single nurse cell nuclei of egg chambers expressing GFP-tagged Zucchini (left) or Papi (right) and stained for mitochondria (immuno-staining of ATP synthase). **f**, Shown are mappings of piRNAs (5' ends only; red, sense; black, antisense) from ovaries of indicated genotype to a second reporter

construct (as in Fig. 1d). Values are normalized to 1 million sequenced miRNA reads. **g**, Scatter plot displaying miRNA-normalized piRNA levels mapping to 63 germline-dominant TEs in Zucchini-depleted versus Zucchini/Papi-depleted ovaries. **h**, Shown are confocal sections of egg chambers expressing indicated piRNA biogenesis reporters (GFP-fluorescence; DNA stained with DAPI) in wild-type (top) or Zucchini and Nibbler-depleted ovaries (middle and bottom). Top, reporter with no target site; middle, reporter used in Fig. 1d; bottom, reporter used in Extended Data Fig. 2f. **i**, Ovary lysates expressing N-terminally Flag-tagged wild-type Nibbler was immunoprecipitated (IPed) using M2 magnetic beads. Wild-type ovary lysates were used as a control. Red colour in the blot indicates a saturated signal. ATP synthase 5A (ATP syn) serves as loading control. Ovary volume indicates the amount of loaded lysate/IP fraction. **j**, Eluates from IP were blotted with indicated antibodies. Piwi is slightly enriched in the Flag-Nibbler IP fraction, while there is no detectable enrichment of Aub or Ago3 in the IP fraction.



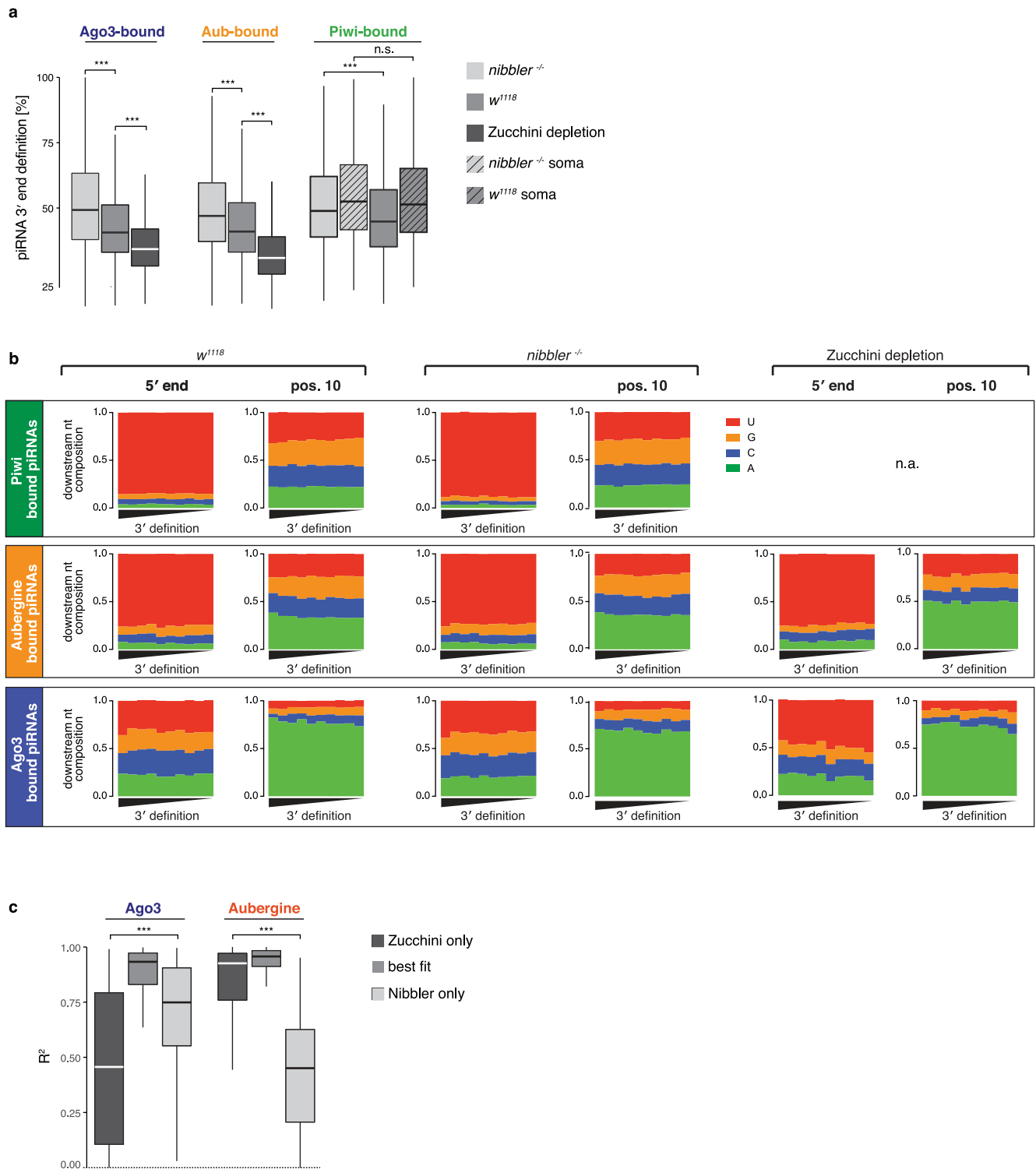
Extended Data Figure 3 | Molecular characterization of the piRNA pathway in *nibbler* and *papi* mutant flies. **a**, *papi* gene locus indicating the position of the Cas9-induced frameshift allele. **b**, *nibbler* gene locus indicating the position of the Cas9-induced frameshift allele. **c**, Western blot analysis showing the loss of detectable Nibbler protein in *nibbler^{-/-}* ovaries. ATP-synthase 5A antibody is used as loading control. Loaded amounts of ovary lysates are indicated. **d**, Northern blot analysis of the Nibbler-substrate miR-34 comparing small RNAs obtained from ovaries of *w¹¹¹⁸* or *nibbler^{-/-}* flies. The GFP-Nibbler rescue transgene (used in Fig. 1f) restores miR-34 processing. **e**, **f**, Immunostainings (**e**) and western blot analysis (**f**) of Piwi, Aub and Ago3 in *w¹¹¹⁸* or *nibbler^{-/-}* ovaries,

showing that localization and expression of the three PIWI-clade proteins are unperturbed (arrow heads; scale bars, 10 μ m). ATP synthase 5A (ATP syn) served as loading control. **g**, Scatter plot displaying the steady-state RNA level of TEs in indicated genetic background (only TEs with RPKM > 1 in either background; $n = 40$). **h**, Bar chart displaying TE mapping piRNA levels in *w¹¹¹⁸* or *nibbler^{-/-}* ovaries (values normalized to 1 million sequenced miRNA reads). **i**, **j**, Length profiles of TE mapping small RNA reads obtained from ovaries of indicated genotypes. Shown are all ovarian small RNAs (**i**) or Piwi-bound piRNAs defined as soma-enriched (**j**; see Methods). Displayed are fractions of reads of indicated length as a percentage (mean lengths are indicated below).



Extended Data Figure 4 | A small RNA library cloning approach that allows the recovery of longer piRNA species. *Drosophila* total RNA contains large amounts of the 30-nt long 2S rRNA. Previous cloning approaches therefore typically restrict small RNA cloning to the 18–29 nt window by cutting these small RNA populations from a gel. We used a previously published 2S rRNA depletion method³⁵, followed by extracting

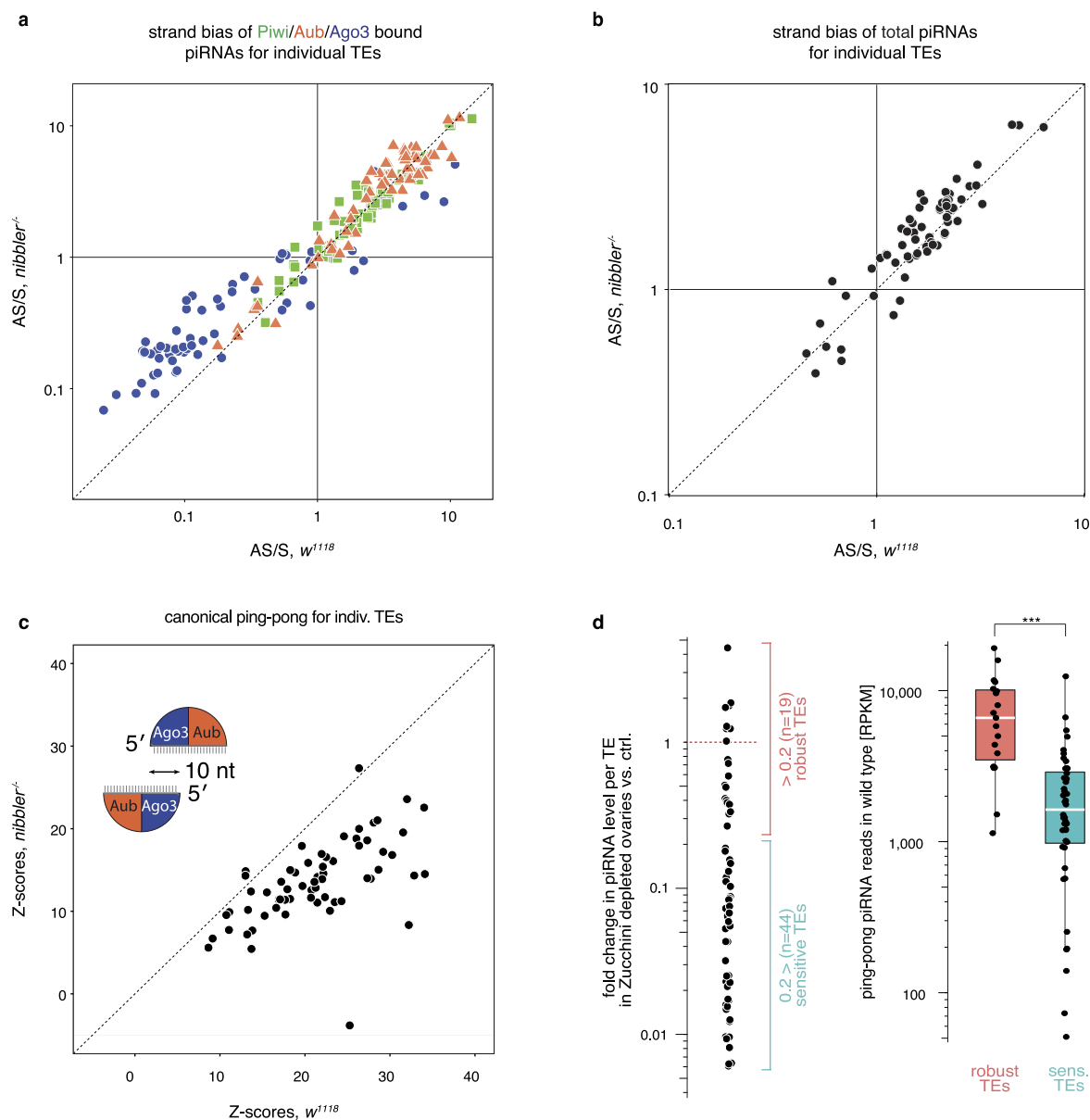
small RNAs ranging from 18–40 nt in length for library preparation. Shown are size distributions of TE mapping small RNAs (obtained from *w¹¹¹⁸* ovaries) comparing the standard small RNA cloning protocol (left) and the protocol using total RNA depleted of 2S rRNA (middle; see Methods). An overlay of the longer reads (>27 nt) is displayed to the right.



Extended Data Figure 5 | Zucchini and Nibbler generate Aub- and Ago3-bound piRNA 3' ends independently of piRNA 5'-end formation.

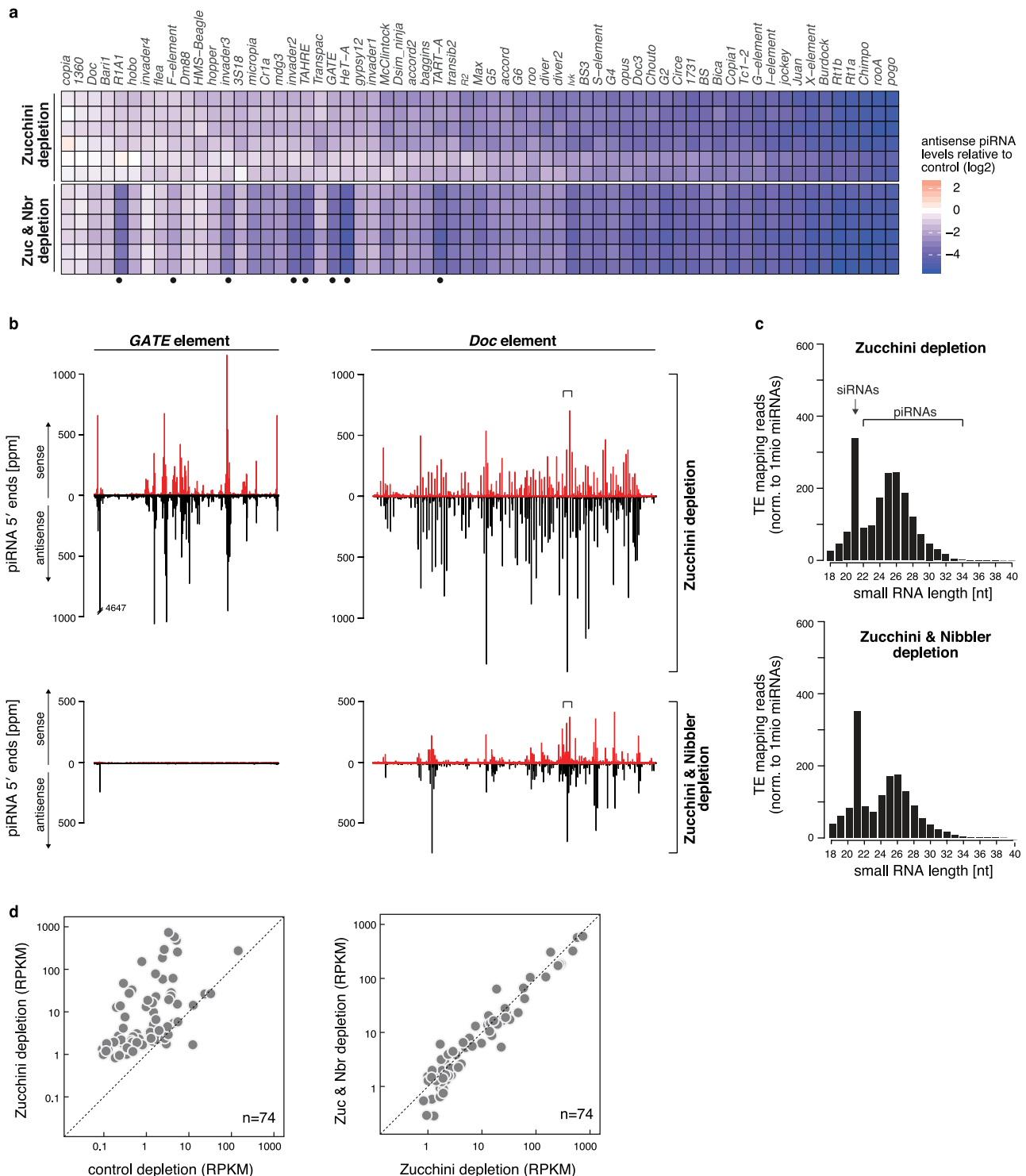
a, Box plots (** $P < 0.001$ by two-sided t -test) showing the 3'-end definition (see Methods) of Ago3-, Aub-, and Piwi-bound piRNAs isolated from ovaries of the indicated genotypes. Soma-enriched Piwi-bound piRNAs (see Methods) are shown in boxes with diagonal lines. **b**, Stacked bar plots displaying the nucleotide composition at the 5' end or position 10 of piRNAs bound to Aub-, Ago3-, or Piwi (isolated from ovaries of indicated genotypes). The plots show the composition within ten equally

sized bins, sorted for their 3' end precision index (see Methods). **c**, Box plots (** $P < 0.001$ by Wilcoxon rank-sum test) showing the R^2 values of the comparison between the 3'-end profiles of Ago3- or Aub-bound piRNAs from *w¹¹¹⁸* ovaries and those of *nibbler*^{-/-} ovaries (Zucchini only), those of ovaries depleted of Zucchini (Nibbler only), and those of the calculated composite that provides the highest R^2 (best fit). Midlines in **a** and **c** indicate the median value, box ranges from the first to the third quartile, whiskers are $1.5 \times$ the interquartile range



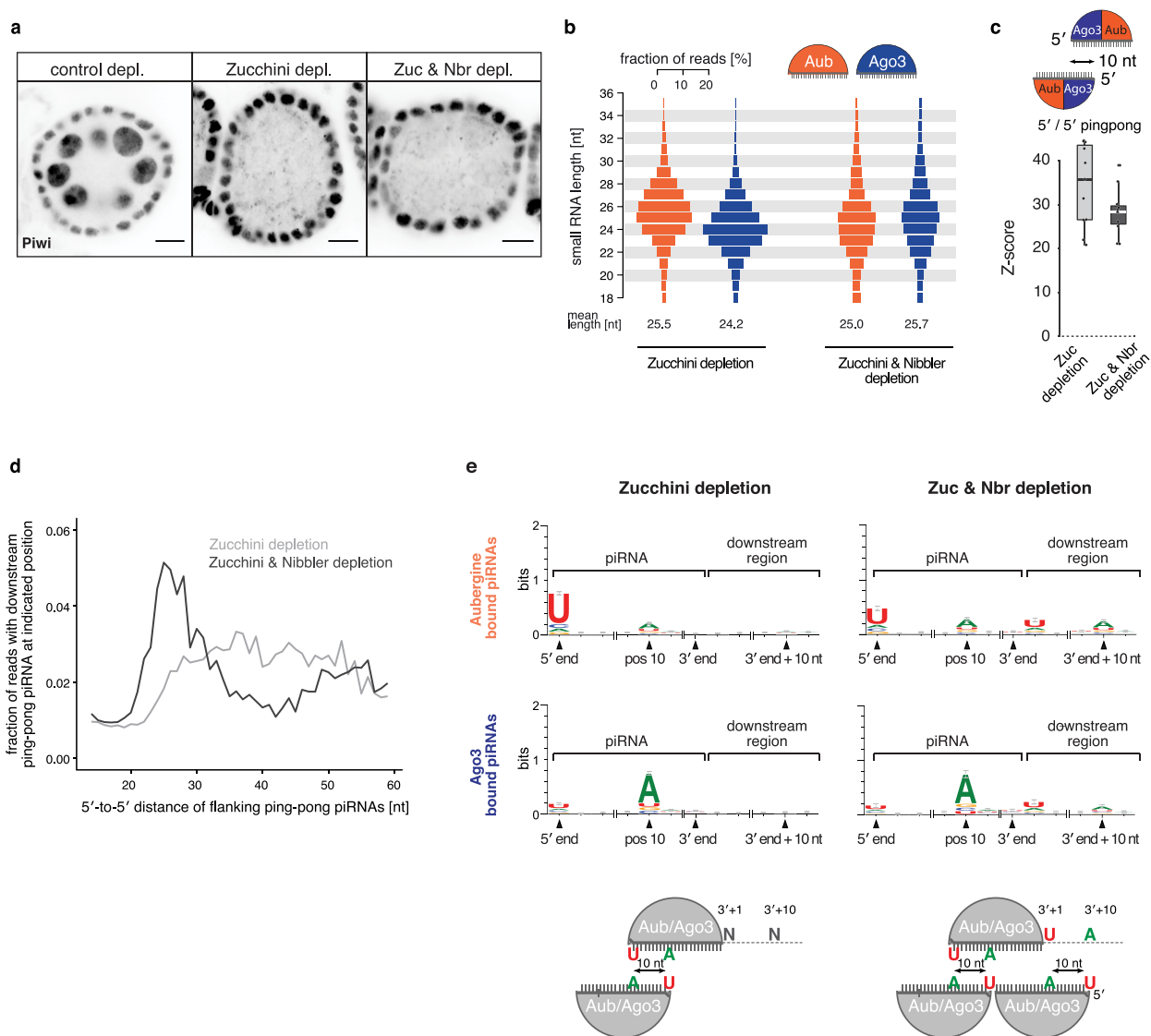
Extended Data Figure 6 | Ago3 incorporates more TE antisense piRNAs in *nibbler* mutant flies. **a, b,** Scatter plots displaying the strand bias (antisense divided by sense) of piRNA populations in *w*¹¹¹⁸ versus *nibbler*^{-/-} ovaries. **a,** piRNAs bound to Piwi, Aub, or Ago3. **b,** piRNAs from total ovarian RNA. **c,** Z-scores of 5'/5' ping-pong levels per TE (63 germline dominant TEs) in *w*¹¹¹⁸ versus *nibbler*^{-/-} ovaries.

d, Grouping of TEs (63 germline dominant TEs) based on fold change in piRNA levels between Zucchini-depleted and control ovaries (left). Box plot indicates Aub/Ago3-bound piRNA levels in wild-type ovaries for defined TE groups (midline indicates the median value, box the first and third quartiles, whiskers are 1.5× the interquartile range; ****P* < 0.001 after Wilcoxon rank-sum test).



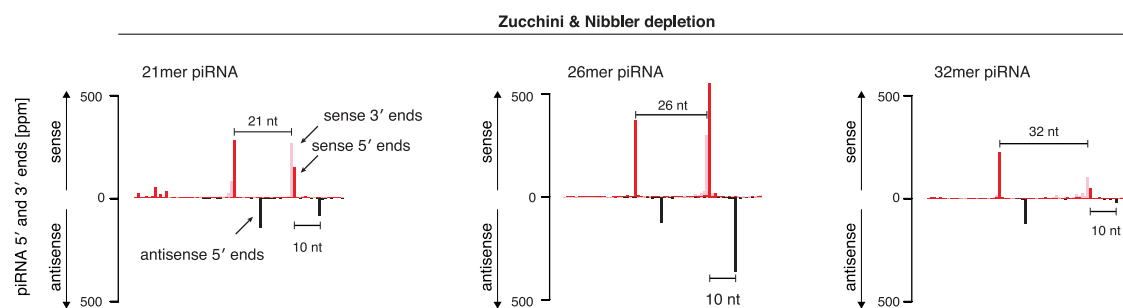
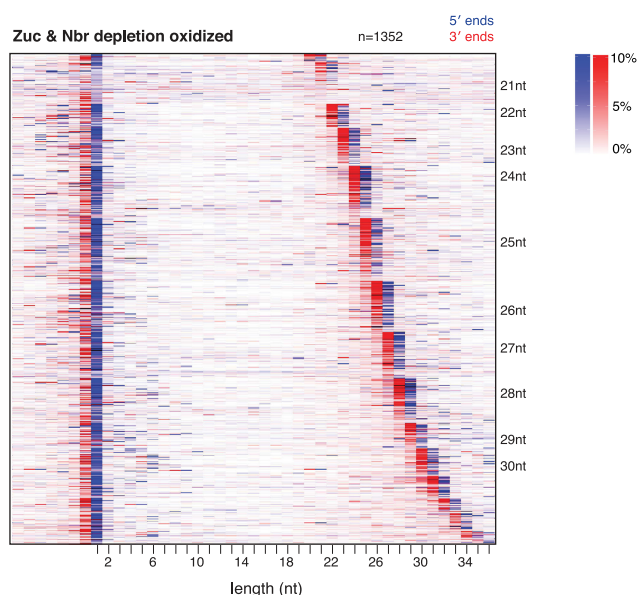
Extended Data Figure 7 | piRNAs are abundantly produced in Zucchini/Nibbler double-depleted ovaries. **a**, Heat map displaying fold changes in antisense piRNA levels per TE ($n = 63$) in indicated genotypes versus control (six biological replicates). Dots mark TEs, which lose piRNAs only in Zucchini/Nibbler-depleted ovaries. **b**, Mappings (5' ends only; red, sense; black, antisense) of piRNAs onto the *GATE* and *Doc* TE consensus sequences. The plots at the top are from a piRNA library obtained from Zucchini-depleted ovaries, and the ones at the bottom are from a piRNA library obtained from Zucchini/Nibbler double-depleted ovaries. The

~100-bp window of *Doc* that is detailed in Fig. 4a is depicted by brackets. **c**, Size distributions of TE-mapping small RNAs obtained from total small RNA libraries from Zucchini-depleted (top) or Zucchini/Nibbler-depleted (bottom) ovaries. Shown are the average values from six biological replicates (reads were normalized to 1 million sequenced miRNA reads). **d**, Scatter plots displaying the steady-state RNA level of TEs in indicated genetic background (only TEs with RPKM > 1 in either background are shown; $n = 74$).



Extended Data Figure 8 | Precise coupling of neighbouring ping-pong piRNAs in Zucchini/Nibbler-depleted ovaries. **a**, Immunofluorescence images (confocal sections) of egg chambers of indicated genotype stained for endogenous Piwi protein (scale bars, 10 μ m). Note that the shRNA-mediated knockdown is specific for germline cells. Somatic follicle cells that also express Piwi therefore serve as control. **b**, Length profiles (fractions of reads per indicated length as a percentage) of TE-mapping piRNAs bound to Aub/Ago3 in indicated genotypes. **c**, Box plot (midline indicates the median value, box the first and third quartiles, whiskers are 1.5 \times the interquartile range) displaying the distribution of Z-scores for canonical 5'/5' ping-pong of piRNA populations isolated from ovaries of indicated genotypes. The analysis is restricted to the 11 germline enriched TEs that maintain piRNA production in Zucchini/Nibbler-depleted

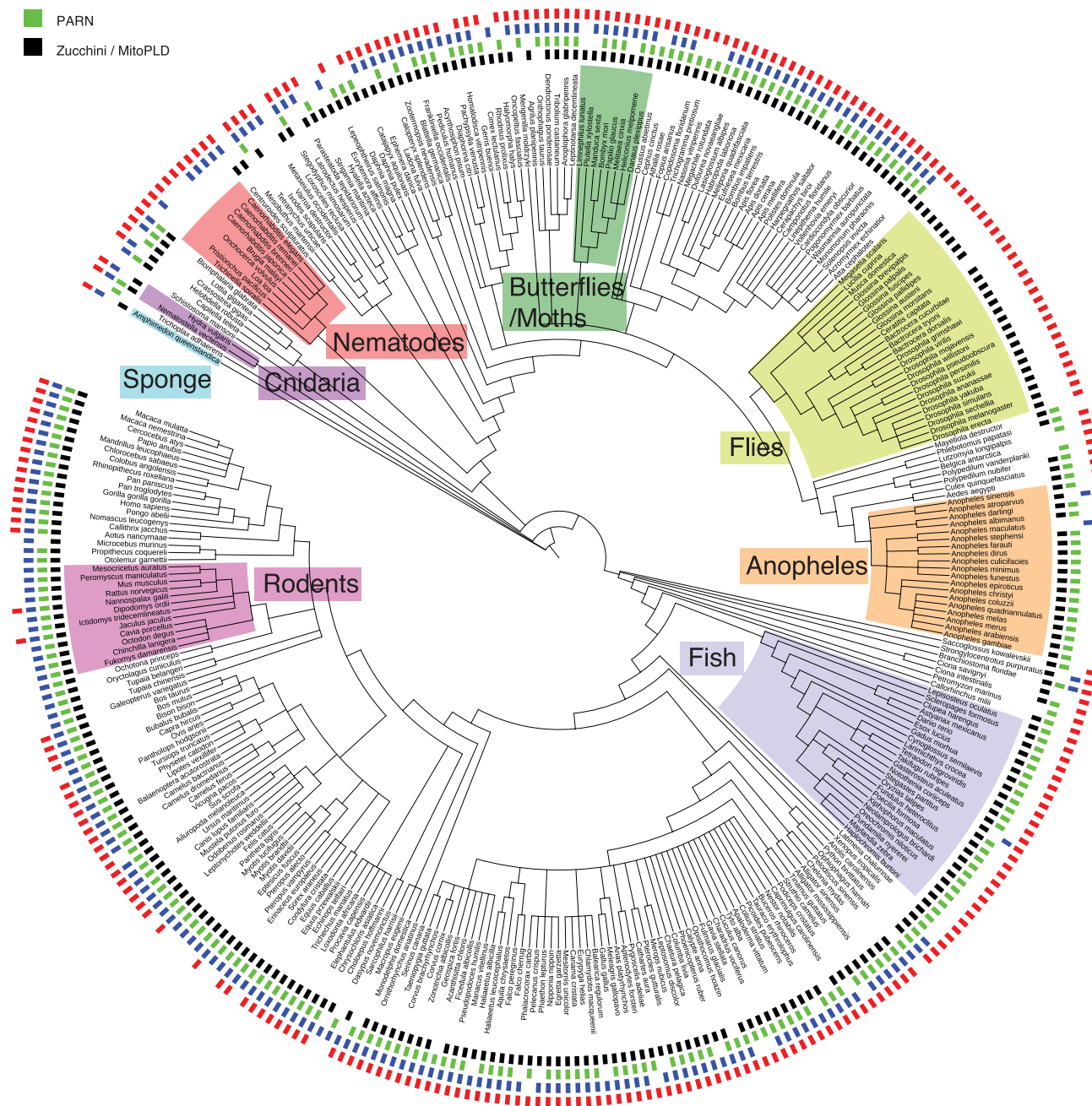
ovaries (see Methods). Midline indicates the median value, box ranges from the first to the third quartile, whiskers are 1.5 \times the interquartile range. **d**, Histogram showing the frequencies of a cloned ping-pong piRNA 5' end downstream of a responder piRNA 5' end at a certain distance on the same strand. **e**, Sequence logos displaying the nucleotide composition within and downstream of Aub- and Ago3-bound piRNAs cloned from ovaries of indicated genotypes. 5'-end position and position 10 are measured from the piRNA 5' end. 3'-end position is the dominant 3' end of a certain piRNA 5' species. Downstream positions are anchored by the dominant 3'-end position. Nucleotide signatures with respect to the position of sense and antisense slicer piRNAs are depicted in the cartoons below.

a**b**

Extended Data Figure 9 | piRNAs generated by slicer/slicer are methylated. **a**, Shown are the mappings of those piRNAs (5' ends in black; 3' ends in grey) that were probed by northern blots in Fig. 4f. The 5' ends of antisense piRNAs (red) map precisely 10 nt downstream of the predicted

slicer sites. Mappings were normalized to 1 million sequenced miRNA reads. **b**, An oxidized library control of the heat map shown in Fig. 4d, showing the precise coupling of piRNA 5' ends (blue) and 3' ends (red) in Zucchini/Nibbler-depleted ovaries.

■ Nibbler / Mut7
■ PNLD1
■ PARN
■ Zucchini / MitoPLD



Extended Data Figure 10 | Conservation of Zucchini/MitoPLD, PARN, PNLD1, and Nibbler/Mut-7 across metazoa. 328 metazoan species are displayed in a phylogenetic tree (generated using iTOL). The presence of indicated orthologues in each species is marked (black, Zucchini/MitoPLD; green, PARN; blue, PNLD1; red, Nibbler/Mut-7). Taxonomic groups mentioned in the text are highlighted (for details see Methods).

An oxidative *N*-demethylase reveals PAS transition from ubiquitous sensor to enzyme

Mary Ortmayr¹, Pierre Lafite^{1†}, Binuraj R. K. Menon¹, Tewes Tralau^{1†}, Karl Fisher¹, Lukas Denkhäus¹, Nigel S. Scrutton¹, Stephen E. J. Rigby¹, Andrew W. Munro¹, Sam Hay¹ & David Leys¹

The universal Per-ARNT-Sim (PAS) domain functions as a signal transduction module involved in sensing diverse stimuli such as small molecules, light, redox state and gases^{1,2}. The highly evolvable PAS scaffold can bind a broad range of ligands, including haem, flavins and metal ions. However, although these ligands can support catalytic activity, to our knowledge no enzymatic PAS domain has been found. Here we report characterization of the first PAS enzyme: a haem-dependent oxidative *N*-demethylase. Unrelated to other amine oxidases, this enzyme contains haem, flavin mononucleotide, 2Fe-2S and tetrahydrofolate cofactors, and specifically catalyses the NADPH-dependent oxidation of dimethylamine. The structure of the α subunit reveals that it is a haem-binding PAS domain, similar in structure to PAS gas sensors³. The dimethylamine substrate forms part of a highly polarized oxygen-binding site, and directly assists oxygen activation by acting as both an electron and proton donor. Our data reveal that the ubiquitous PAS domain can make the transition from sensor to enzyme, suggesting that the PAS scaffold can support the development of artificial enzymes.

The PAS domain has been found in all kingdoms of life, functioning as a signal transducer that couples ligand binding to a cellular signalling response^{1,2}. The highly divergent PAS family couples a remarkable plasticity in binding ligands to an equally wide variety of sensory stimuli. The typical PAS fold consists of a single antiparallel, five stranded β -sheet with intervening α -helices that create the ligand-binding pocket on the β -sheet surface. Despite possessing structural similarity, PAS domains often have low sequence similarity, with no universally conserved residues. This suggests a high adaptability, which has been experimentally verified for photoactive yellow protein⁴. Despite PAS domains binding cofactors with inherent chemical reactivity, no enzymatic activity has been associated with the PAS family. By contrast, other folds have made the transition to enzyme, such as globins^{5,6}. Given the remarkable plasticity and universal occurrence of PAS, this might suggest an inherent limitation to developing enzymatic activity. Understanding of such limit(s) is key to the emerging field of artificial enzyme development, which often makes use of ligand-binding scaffolds⁷.

We show that an orphan enzyme previously described as a secondary amine mono-oxygenase^{8,9}, but here referred to as haem-dependent oxidative *N*-demethylase (HODM), contains a haem-dependent PAS-scaffold that serves as the site of the oxidative *N*-demethylase activity. HODM functions in bacterial methylated amine catabolism, linking alkylamine oxidation to the tetrahydrofolate C1 pool¹⁰. Methylated amines are formed through various processes¹¹ and the trimethylamine degradation pathway has been outlined since the 1970s (Extended Data Fig. 1a). Previous spectroscopic characterization of HODM revealed that it combines features typically associated with haem proteins as diverse as globins, cytochromes P450 and cytochrome *b*₅ (refs 8, 9). The NADPH-dependent dimethylamine (DMA) oxidation reaction catalysed by HODM appears similar to a typical cytochrome P450 *N*-dealkylation reaction¹², although HODM lacks the characteristic

P450 cysteine ligation. Whether the enzyme catalyses a P450-like mono-oxygenation has not been established. Until recently, further studies have been hampered by the lack of sequence information¹³.

We expressed the operon proposed to encode for HODM from the proteobacterium *Pseudomonas mendocina* in *Escherichia coli*. This includes a gene coding for a DUF3445 family representative of unknown function (42 kDa, α), a phthalate dioxygenase reductase-like FMN/2Fe-2S dependent oxidoreductase (39 kDa, β), a glycine cleavage T-protein (34 kDa, γ) and a smaller protein of unknown function (19 kDa, δ subunit) (Fig. 1a). The purified gene products form a heterotetrameric enzyme (Extended Data Fig. 1a) with biochemical and spectral properties similar to the *Aminobacter aminovorans* HODM^{8,9} (Fig. 1).

Electron paramagnetic resonance spectroscopy of anaerobically isolated HODM reveals signals for all redox centres, including the presence of a NO-bound ferrous haem (Fig. 1b–f). HODM catalyses the NADPH-dependent oxidative *N*-demethylation of DMA with a k_{cat} of $9.2 \pm 0.5 \text{ s}^{-1}$ and a $K_{\text{m}}^{\text{DMA}}$ of $24 \pm 6 \mu\text{M}$, resulting in the release of one equivalent of formaldehyde under *in vitro* conditions (that is, no THF present; Fig. 1g). The presence of the γ subunit significantly lowers the DMA-dependent formaldehyde release *in vivo* (Extended Data Fig. 1b). This suggests substrate channelling occurs between the α and γ subunits, as observed for other amine oxidases^{14,15}. Binding of oxygen to reduced HODM results in globin-type spectral changes (Fig. 1h), and the resulting ferrous–oxygen complex is remarkably stable (half-life of approximately 50 min, Extended Data Fig. 1c), probably linked to the high haem redox potential (Extended Data Fig. 1d). DMA binding perturbs the haem spectrum, and DMA affinity is significantly higher for the ferrous–oxygen complex ($K_{\text{d}} = 15 \pm 3 \mu\text{M}$) (Extended Data Fig. 1e, f).

DMA binding to anaerobically purified HODM leads to the loss of superhyperfine coupling to the ¹⁴N atom of the His ligand in the NO-ferrohaem spectrum (Extended Data Fig. 2), probably a consequence of a change in the relative orientation of the NO and histidine ligands to the haem iron^{16,17}. This suggests DMA binding affects the NO ligand, in line with the allosteric effect of oxygen binding on DMA affinity. DMA binding also leads to a change in the *g* values exhibited by the [2Fe-2S]¹⁺ cluster (Extended Data Fig. 2d), suggesting communication between the β and α subunits. These effects only occur with DMA and not with poor substrates (Extended Data Fig. 2e–h).

While NADPH-dependent flavin reduction occurs rapidly ($k = 304 \pm 25 \text{ s}^{-1}$ under anaerobic conditions; Fig. 2a), haem reduction is extremely slow in the absence of both DMA and oxygen ($k = 0.059 \pm 0.005 \text{ min}^{-1}$; Fig. 2b). Haem reduction remains slow in the presence of either oxygen (leading to the ferrous–O₂ complex) or DMA alone (leading to the ferrous–DMA complex), but rapid mixing of the HODM ferrous–oxygen–DMA complex (containing oxidized flavin mononucleotide (FMN)/[2Fe-2S]) with NADPH under anaerobic conditions reveals that the presence of substrate enhances the initial

¹Manchester Institute of Biotechnology, School of Chemistry, 131 Princess Street, University of Manchester, Manchester M1 7DN, UK. [†]Present addresses: University of Orléans, CNRS, ICOA, UMR 7311, F-45067 Orléans, France (P.L.); German Federal Institute for Risk Assessment (BfR), Department of Chemicals and Product Safety, Max-Dohrn Str. 8-10, 10589 Berlin, Germany (T.T.).

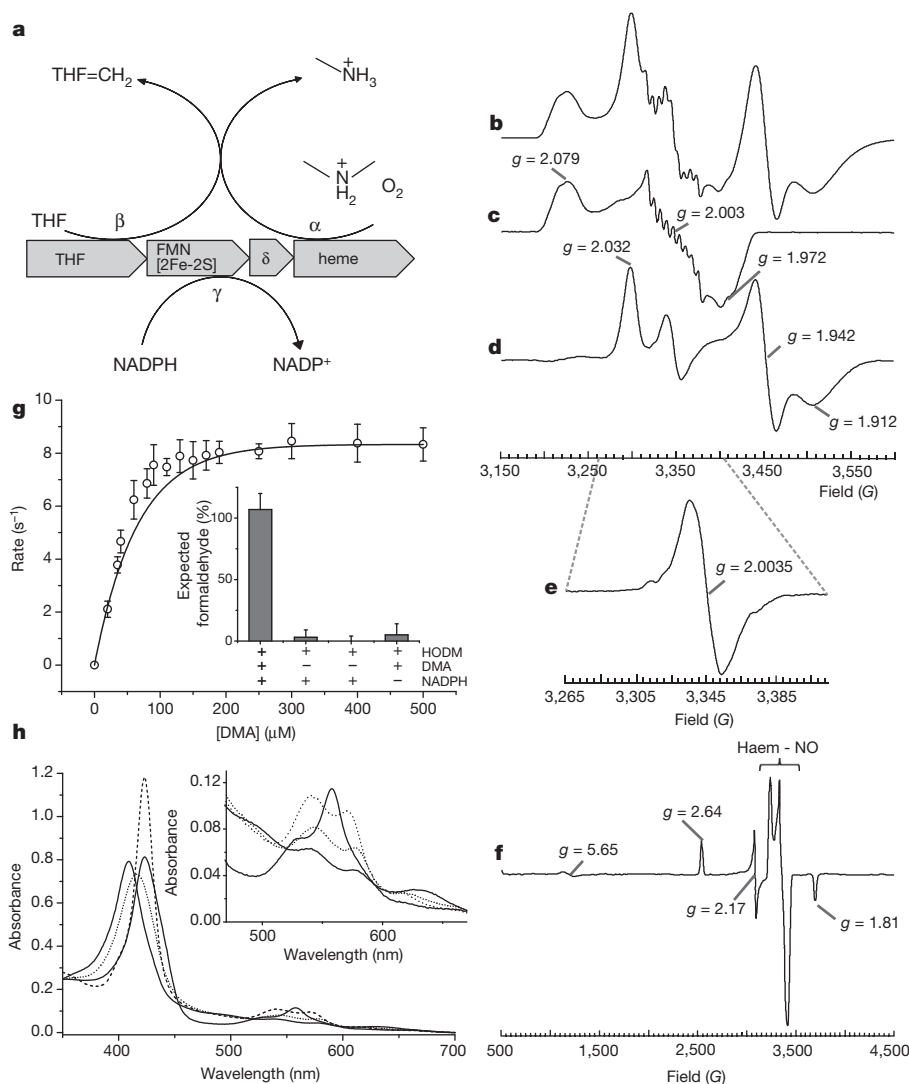


Figure 1 | Biophysical characterization of *P. mendocina* HODM.

a, Schematic overview of the HODM reaction. **b–f**, X-band continuous wave EPR spectra of: anaerobically purified HODM (**b**); HODM exposed to air (**c**); the difference spectrum formed by the subtraction of **c** from **b** showing rhombic $[2\text{Fe-2S}]^{1+}$ and flavin semiquinone contributions³⁰ (**d**); anaerobically purified HODM spectrum around $g = 2.00$ run under conditions designed to emphasize the neutral flavin semiquinone signal at

$g = 2.00$ in **d** (**e**); and potassium ferricyanide oxidized HODM (**f**). **g**, DMA-dependent NADPH oxidation ($k_{\text{cat}} = 9.2 \pm 0.5 \text{ s}^{-1}$ and $K_{\text{M}}^{\text{DMA}} = 24 \pm 6 \mu\text{M}$). The inset reveals stoichiometric conversion of DMA to MA and formaldehyde (errors bars are s.e.m., $n = 3$). **h**, UV-visible absorption features of $8.5 \mu\text{M}$ HODM in the ferric (thick solid line), ferrous (thin solid line), $\text{Fe}^{2+}\text{-O}_2$ (dotted line) and $\text{Fe}^{2+}\text{-CO}$ (dashed line) states.

β reduction rate by $>1,000$ fold, and the ferrous–oxygen–DMA complex decay rate by 2.5-fold (Fig. 2c–e). A rate enhancement for the β subunit reduction and intramolecular electron transfer to the ferrous–oxygen–DMA (as opposed to the ferrous–oxygen) complex is probably linked to the observed effects of DMA binding on the $[2\text{Fe-2S}]^{1+}$ cluster. However, the ferric to ferrous haem iron transition under anaerobic conditions remains around 100 times slower (both with and without DMA) compared to the k_{cat} of $9.2 \pm 0.5 \text{ s}^{-1}$.

Following the aerobic mixing of HODM with excess DMA and NADPH, steady-state conditions are established, during which the enzyme remains predominantly in the ferric–DMA-bound state (Fig. 2e). This finding suggests that intermolecular electron transfer is rate-limiting, supported by the fact that no significant kinetic isotope effect (KIE) on the NADPH oxidation rates could be observed using perdeuterated DMA or D_2O . The mechanism that underpins the rate enhancement observed only in the presence of both DMA and O_2 remains unclear at this stage.

As observed for holo-HODM, reduction of the isolated α subunit is a slow process. Mixing of the ferrous– O_2 –DMA α subunit with NADH and phenazine methosulfate under anaerobic conditions leads

to accumulation of the ferrous–DMA enzyme via transient formation of the ferric–DMA species (Fig. 2f). Low nM (sub-stoichiometric) levels of formaldehyde can be detected following this experiment, confirming that the α subunit is catalytically active. Although certain haem enzymes can make use of peroxide to bypass the initial oxygen reduction step, no spectroscopic evidence for higher Fe oxidation species could be obtained in the presence of DMA and peroxide, nor was formaldehyde formation observed under these conditions (Extended Data Fig. 3).

Unfortunately, we were unable to determine the structure of the HODM holoenzyme. However, the isolated α subunit responsible for haem, oxygen and DMA binding has spectral and ligand binding properties nearly identical to those of the holoenzyme (Extended Data Fig. 4), is catalytically competent and amenable to crystallization. The 1.7 Å resolution α subunit structure reveals that it consists of an N- and C-terminally extended PAS haem binding domain (Fig. 3), most similar in structure to the Aer2 PAS domain, a gas-sensing receptor protein controlling chemotaxis in *Pseudomonas aeruginosa*¹⁸ (PDB code 4HI4; 2.25 Å r.m.s.d. over 85 C α atoms, Z score 5.6, 12% homology; Fig. 3b). The HODM PAS domain (residues 145–231) contains the canonical

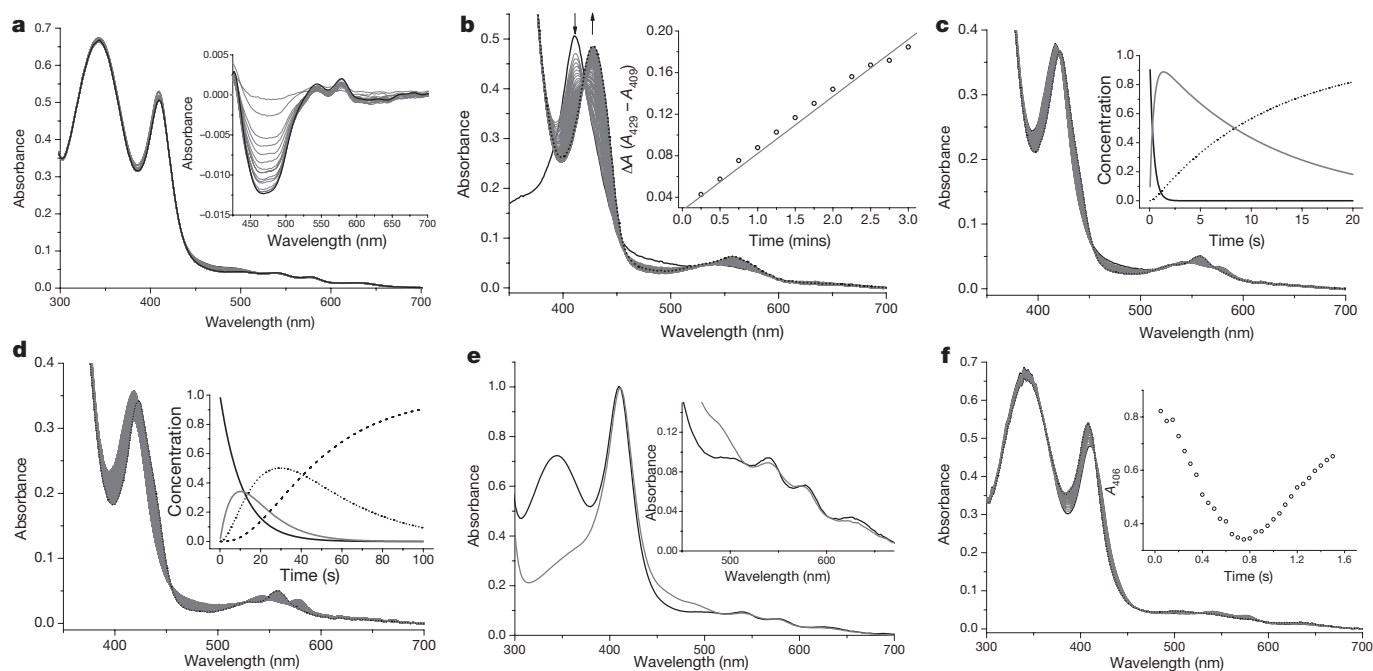


Figure 2 | NADPH-dependent oxidation of DMA by HODM. **a**, UV-visible spectral evolution (over 1 s) from the anaerobic mixing of ferric HODM (5.3 μM) with NADPH. Difference spectra (inset) reveal the γ subunit flavin is rapidly reduced. No haem reduction is observed. **b**, As in **a**, but monitored over 3 min. Arrows indicate absorbance changes from ferric (409 nm, thick line) to ferrous (429 nm, thick line). The $A_{429} - A_{409}$ signal (inset) is fitted to an $a \rightarrow b$ model; $k = 0.059 \pm 0.005 \text{ min}^{-1}$ (errors bars are s.e.m., $n = 3$). **c**, UV-visible spectral evolution (over 20 s) from the anaerobic mixing of ferrous-oxy HODM (4 μM) with 200 μM NADPH (in the presence of 2 mM DMA). An $a \rightarrow b \rightarrow c$ model was fitted with $k_1 = 2.58 \pm 0.05 \text{ s}^{-1}$ for ferrous-oxy-DMA (black line) to ferric (grey line) and $k_2 = 0.087 \pm 0.007 \text{ s}^{-1}$ for the ferric-DMA to ferrous-DMA (dotted line). Similar experiments over a 1 s period revealed these steps are preceded by flavin reduction ($k = 149 \pm 20 \text{ s}^{-1}$). **d**, UV-visible spectral

evolution (over 100 s) from the anaerobic mixing of ferrous-oxy HODM (4 μM) with 200 μM NADPH (in the absence of DMA). Fitted to an $a \rightarrow b \rightarrow c \rightarrow d$ model with $k_1 = 0.09 \pm 0.02 \text{ s}^{-1}$ (initial β subunit reduction; black line to grey line), $k_2 = 1.0 \pm 0.2 \text{ s}^{-1}$ for the ferrous-oxy-DMA to ferric (dotted line) and $k_3 = 0.03 \pm 0.01 \text{ s}^{-1}$ for the ferric to ferrous (dashed line) transition. **e**, An overlay of the HODM spectrum under steady-state conditions (enzyme in the presence of excess NADPH, O_2 and DMA; black line) with the ferric DMA-bound resting state (grey line). **f**, UV-visible spectral evolution (over 1.6 s) from the anaerobic mixing of ferrous-oxy HODM α subunit (7.4 μM) with NADH in the presence of 75 μM phenazine methosulfate and 2 mM DMA. Fitted to an $a \rightarrow b \rightarrow c$ model with $k_1 = 0.24 \pm 0.07 \text{ s}^{-1}$ (ferrous-oxy-DMA to ferric) and $k_2 = 0.09 \pm 0.02 \text{ s}^{-1}$ (ferric-DMA to ferrous-DMA).

A β , B β , G β , H β , and I β PAS β -sheet as well as the core D α , E α and F α elements. Similar to Aer2, HODM includes an extended H β –I β loop and a shifted F α helix not seen in other haem-binding PAS domains. The HODM E η 1 His194 acts as the proximal haem ligand, comparable to Aer2 E η 1 His234. However, HODM lacks the core PAS-domain C α helix that encompasses the Aer2 haem pocket, and features extended β -strands as well as a longer G β –H β loop. The N- and C-terminal extensions to the central HODM PAS domain cover the PAS β -sheet on the side opposite to the haem binding pocket.

Although the stoichiometry of the overall reaction catalysed by HODM bears some similarity to the cytochrome P450 enzymes¹², HODM does not contain a Cys haem-thiolate ligand that assists with oxygen–oxygen heterolytic bond cleavage to form compound I (the so-called ‘push’ effect¹⁹). By contrast, many haem peroxidases also contain a His ligand and have been shown to contain a highly polarized distal active site that achieves the same feat (referred to as the ‘pull’ mechanism¹⁹). We determined the structures of HODM α in complex with DMA (to 1.75 Å resolution, Fig. 3d) and with both DMA and NO (a mimic for the DMA– O_2 complex; to 1.9 Å resolution, Fig. 3e). DMA binding displaces two water molecules and occurs in close proximity of the distal haem ligand (Fig. 3d). Direct electrostatic interactions between the DMA substrate and the enzyme are made with the conserved H β Glu266, in addition to van der Waals interactions between the substrate methyl groups, Trp180 and the haem porphyrin. NO binding to the haem iron occurs at a distance of 1.95 Å (Fig. 3e). The oxygen atom of NO (mimicking the distal β oxygen of O_2) is within hydrogen bonding distance of the conserved G β Arg224 and H β Gln268, as well as the DMA amine group (Fig. 3e). This is

consistent with the approximately 90-fold higher DMA affinity of the ferrous-oxy complex compared with ferrous HODM. Substitution of Trp180, Arg224 or Glu266 all lead to enzyme inactivation (data not shown). Although oxygen binding is detected for W180A, E266Q and R224A variants, no DMA binding is detected for E266Q, while direct haem ligation by DMA occurs in the R224A variant (Extended Data Fig. 4c–e).

The structural arrangement of the HODM haem subunit active site shares the polarized nature of the haem peroxidases²⁰: the conserved Arg224, the amine substrate (presumably charged when bound to conserved H β Glu266) and the conserved Gln268 each contribute a (partial) positive charge to the distal oxygen. On the other hand, the HODM proximal ligand environment is similar to that observed in globins, and a peroxidase-like interaction with an Asp to enhance electron-donating character is not observed. The latter is consistent with the stable O_2 complex and the positive redox potential. A key question is whether HODM achieves oxidative *N*-demethylation through species with higher Fe oxidation states, such as the highly reactive compound I (as typified by cytochrome P450 enzymes¹²) or compound II (as occurs in peroxidases²⁰).

We propose that, following DMA binding, the β subunit reduces the ferric haem iron to ferrous using NADPH, leading to dioxygen binding and formation of the ferrous-oxy/ferric-superoxy form (Fig. 4). Transfer of the second electron and formation of the ferric hydroperoxide complex (compound 0) is probably coupled to proton transfer from the DMA substrate. This particular step explains the strict requirement of this enzyme for amine substrates. Following deprotonation of the substrate concomitant with compound 0 formation, Arg224 is likely

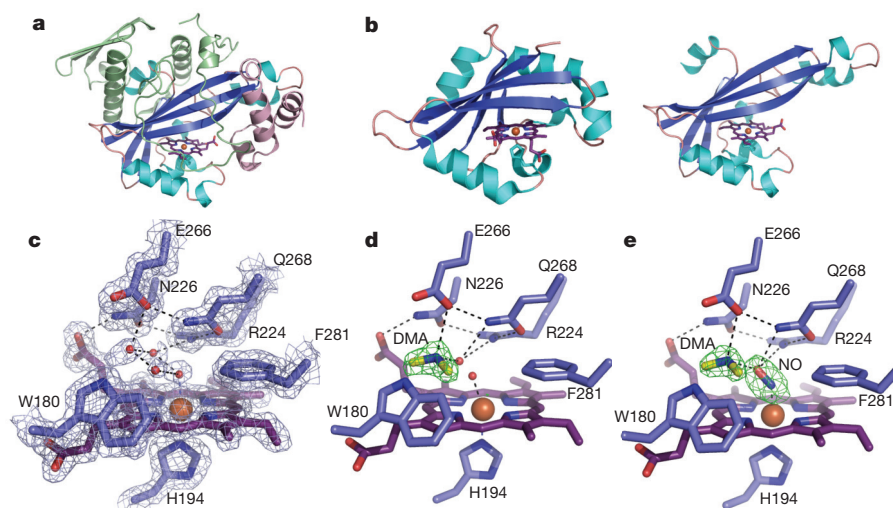


Figure 3 | Crystal structure of the *P. mendocina* HODM α subunit.

a, Cartoon representation of HODM α , with the N-terminal and the C-terminal extensions coloured in green and pink respectively. The central PAS domain is coloured according to secondary structure. The haem cofactor is shown in atom coloured sticks. **b**, A side-by-side comparison of the HODM PAS domain (left panel) with the PAS oxygen sensor Aer2 (right panel). Colour coding as in **a**. **c**, A view of the HODM α subunit active site in the ferric state, with key residues shown in atom coloured

sticks. Sequence alignment reveals H194, R224 and F281 are strictly conserved, while other positions contain conservative mutations: W or F occur at 180, E or D at 266 and Q or E at 268. The $2F_o - F_c$ electron density map is contoured at 1σ (blue mesh). **d**, HODM α active site in complex with the DMA substrate. The $F_o - F_c$ omit density corresponding to DMA is contoured at 4σ (green mesh). **e**, HODM α active site in complex with the DMA substrate and NO. The $F_o - F_c$ omit density corresponding to DMA and NO is contoured at 4σ (green mesh).

to act as the acid required to catalyse O–O scission. Oxidation of the substrate and O–O scission can, in principle, occur in a variety of ways (Extended Data Fig. 5), possibly including Fe(V) and/or Fe(IV) oxidation states. Both mono-oxygenase (the direct incorporation of oxygen, leading to hydroxymethylamine) as well as amine oxidase activity (leading to an imine product) appear possible, but are not readily distinguishable as both products rapidly convert to formaldehyde and mono-methylamine.

While ample spectroscopic evidence has been presented for such reaction mechanisms in other haem-dependent enzymes^{19–21}, we have found no evidence of either compound I or II during turnover or rapid mixing experiments with HODM. Similarly, X-ray-induced photoreduction of ferrous–O₂–DMA crystals at 100 K led to formation of the ferrous species with no indication of higher Fe oxidation states (Extended Data Fig. 6). Our inability to observe transient, higher haem Fe oxidation states does not exclude the possibility that HODM operates via such species. However, the direct alignment of the substrate *N*-methyl group with bound oxygen also allows for direct hydride transfer from the methyl group to the compound 0 proximal oxygen. Coupled with heterolytic O–O scission and protonation of the distal oxygen by Arg 224, this mechanism does not require formation of a higher Fe oxidation state (Fig. 4). A direct hydride transfer mechanism has been proposed for the nitric oxide reductase P450nor²², while the direct addition of Cl[–] to Fe(V)=O in chloroperoxidases has some similarity²³.

Alternatively, proton transfer from Arg224 could be coupled to single-electron transfer from the substrate, creating compound II and a substrate amine radical. In contrast to the hydride transfer mechanism, the radical species can result in either the *N*-oxide or iminium product. It is interesting to note that, unlike DMA, methylamine (MA) is oxidized without the formation of significant levels of formaldehyde (Extended Data Fig. 7). This suggests that HODM is unable to form the MA-derived imine product. As the MA-derived reaction intermediate(s) are similar in energy to their DMA-derived counterparts (Extended Data Fig. 8), this suggests incorrect positioning of the single methyl group of MA, leading to formation of the *N*-oxide product. Distinct active and inactive methyl binding pockets have been implicated in the sequential *N*-demethylation process catalysed by cytochrome P450 2B4²⁴.

Our data confirm that the PAS fold can accommodate enzymatic activity, which suggests that other PAS enzymes are likely to exist. Furthermore, it points to the fact that existing PAS domains can be evolved to catalyse reactions: engineered variants of a PAS oxygen sensor have been shown to display haem oxygenase activity²⁵. *N*-dealkylation is a widespread reaction, catalysed by cytochrome P450 enzymes¹², FAD/quinone amine oxidases^{26,27} and α -ketoglutarate-dependent Fe oxygenases²⁸. HODM represents a distinct catalyst, which we propose acts as an oxidase. Unlike the more widespread flavin amine oxidases, HODM oxidizes both NADPH and the amine substrate to convert oxygen to water (as opposed to peroxide). DMA oxidation sensitive to CO has been shown to occur in yeasts²⁹, and this new type of amine oxidase also appears in eukaryotes, with HODM α homologous genes present in a range of genomes (Extended Data Fig. 9).

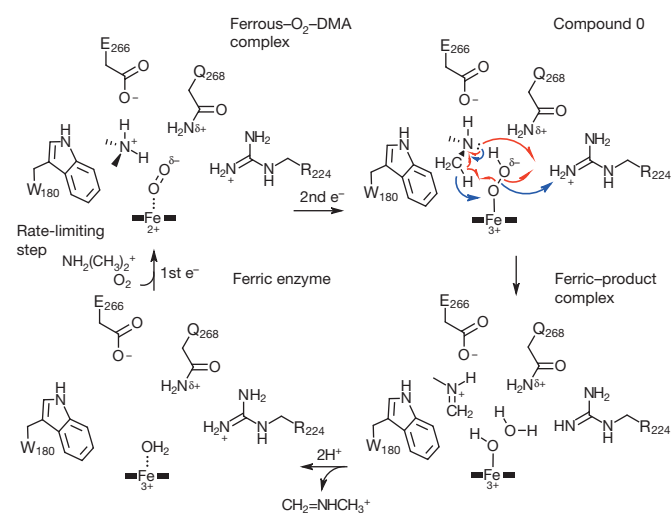


Figure 4 | HODM is a haem-dependent amine oxidase. Following formation of compound 0, DMA oxidation can occur through a radical mechanism (red arrows) involving transient formation of the Fe(IV) oxidation state (Compound II) or directly through H-transfer (blue arrows), bypassing the need for higher Fe oxidation states.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 15 March; accepted 15 October 2016.

Published online 16 November 2016.

1. Henry, J. T. & Crosson, S. Ligand-binding PAS domains in a genomic, cellular, and structural context. *Annu. Rev. Microbiol.* **65**, 261–286 (2011).
2. Möglich, A., Ayers, R. A. & Moffat, K. Structure and signaling mechanism of Per-ARNT-Sim domains. *Structure* **17**, 1282–1294 (2009).
3. Gilles-Gonzalez, M. A. & Gonzalez, G. Heme-based sensors: defining characteristics, recent developments, and regulatory hypotheses. *J. Inorg. Biochem.* **99**, 1–22 (2005).
4. Philip, A. F., Kumauchi, M. & Hoff, W. D. Robustness and evolvability in the functional anatomy of a PER-ARNT-SIM (PAS) domain. *Proc. Natl Acad. Sci. USA* **107**, 17986–17991 (2010).
5. Gardner, P. R. Nitric oxide dioxygenase function and mechanism of flavohemoglobin, hemoglobin, myoglobin and their associated reductases. *J. Inorg. Biochem.* **99**, 247–266 (2005).
6. Sun, S. *et al.* Influence of heme environment structure on dioxygen affinity for the dual function Amphitrite ornata hemoglobin/dehaloperoxidase. Insights into the evolutionary structure-function adaptations. *Arch. Biochem. Biophys.* **545**, 108–115 (2014).
7. Ward, T. R. Artificial metalloenzymes based on the biotin-avidin technology: enantioselective catalysis and beyond. *Acc. Chem. Res.* **44**, 47–57 (2011).
8. Alberta, J. A. & Dawson, J. H. Purification to homogeneity and initial physical characterization of secondary amine monooxygenase. *J. Biol. Chem.* **262**, 11857–11863 (1987).
9. Alberta, J. A., Andersson, L. A. & Dawson, J. H. Spectroscopic characterization of secondary amine mono-oxygenase. Comparison to cytochrome P-450 and myoglobin. *J. Biol. Chem.* **264**, 20467–20473 (1989).
10. Chistoserdova, L. Modularity of methylotrophy, revisited. *Environ. Microbiol.* **13**, 2603–2622 (2011).
11. Colby, J. & Zatman, L. J. Enzymological aspects of the pathways for trimethylamine oxidation and C1 assimilation of obligate methylotrophs and restricted facultative methylotrophs. *Biochem. J.* **148**, 513–520 (1975).
12. Guengerich, F. P. & Munro, A. W. Unusual cytochrome p450 enzymes and reactions. *J. Biol. Chem.* **288**, 17065–17073 (2013).
13. Dziewit, L. *et al.* Genome-guided insight into the methylotrophy of *Paracoccus aminophilus* JCM 7686. *Front. Microbiol.* **6**, 852 (2015).
14. Tralau, T. *et al.* An internal reaction chamber in dimethylglycine oxidase provides efficient protection from exposure to toxic formaldehyde. *J. Biol. Chem.* **284**, 17826–17834 (2009).
15. Singh, H., Arentson, B. W., Becker, D. F. & Tanner, J. J. Structures of the PutA peripheral membrane flavoenzyme reveal a dynamic substrate-channeling tunnel and the quinone-binding site. *Proc. Natl Acad. Sci. USA* **111**, 3389–3394 (2014).
16. LoBrutto, R., Wei, Y.-H., Mascarenhas, R., Scholes, C. P. & King, T. E. Electron nuclear double resonance and electron paramagnetic resonance study on the structure of the NO-ligated heme alpha 3 in cytochrome c oxidase. *J. Biol. Chem.* **258**, 7437–7448 (1983).
17. Morse, R. H. & Chan, S. I. Electron paramagnetic resonance studies of nitrosyl ferrous heme complexes. Determination of an equilibrium between two conformations. *J. Biol. Chem.* **255**, 7876–7882 (1980).
18. Sawai, H. *et al.* Structural basis for oxygen sensing and signal transduction of the heme-based sensor protein Aer2 from *Pseudomonas aeruginosa*. *Chem. Commun. (Camb.)* **48**, 6523–6525 (2012).
19. Sono, M., Roach, M. P., Coulter, E. D. & Dawson, J. H. Heme-Containing Oxygenases. *Chem. Rev.* **96**, 2841–2888 (1996).
20. Matsunaga, I. & Shiro, Y. Peroxide-utilizing biocatalysts: structural and functional diversity of heme-containing enzymes. *Curr. Opin. Chem. Biol.* **8**, 127–132 (2004).
21. Rittle, J. & Green, M. T. Cytochrome P450 compound I: capture, characterization, and C-H bond activation kinetics. *Science* **330**, 933–937 (2010).
22. Daiber, A. *et al.* Isotope effects and intermediates in the reduction of NO by P450_(NO). *J. Inorg. Biochem.* **88**, 343–352 (2002).
23. Hofrichter, M. & Ullrich, R. Heme-thiolate haloperoxidases: versatile biocatalysts with biotechnological and environmental significance. *Appl. Microbiol. Biotechnol.* **71**, 276–288 (2006).
24. Roberts, A. G. *et al.* NMR-derived models of amidopyrine and its metabolites in complexes with rabbit cytochrome P450 2B4 reveal a structural mechanism of sequential N-dealkylation. *Biochemistry* **50**, 2123–2134 (2011).
25. Du, Y. *et al.* Conversion of a heme-based oxygen sensor to a heme oxygenase by hydrogen sulfide: effects of mutations in the heme distal side of a heme-based oxygen sensor phosphodiesterase (Ec DOS). *Biomaterials* **26**, 839–852 (2013).
26. Fitzpatrick, P. F. Oxidation of amines by flavoproteins. *Arch. Biochem. Biophys.* **493**, 13–25 (2010).
27. Brazeau, B. J., Johnson, B. J. & Wilmot, C. M. Copper-containing amine oxidases. Biogenesis and catalysis; a structural perspective. *Arch. Biochem. Biophys.* **428**, 22–31 (2004).
28. Walport, L. J., Hopkinson, R. J. & Schofield, C. J. Mechanisms of human histone and nucleic acid demethylases. *Curr. Opin. Chem. Biol.* **16**, 525–534 (2012).
29. Green, J. & Large, P. J. Oxidation of dimethylamine and trimethylamine in methazotrophic yeasts by microsomal mono-oxygenases sensitive to carbon monoxide. *Biochem. Biophys. Res. Commun.* **113**, 900–907 (1983).
30. Bertrand, P., More, C. & Camensuli, P. Evidence for a magic magnetic configuration between FMN and the [2Fe-2S]⁺ center of phthalate dioxygenase reductase of *Pseudomonas cepacia*. *J. Am. Chem. Soc.* **117**, 1807–1809 (1995).

Acknowledgements This work was supported by BBSRC grant (BBE0170101). We thank the BBSRC/EPSC SYNBIOCHEM Centre (grant BB/M017702/1) for access to analytical equipment. We thank the Diamond Light Source for access to beamlines (proposal number MX8997). D.L. is a Royal Society Wolfson Merit Award holder. N.S.S. is an EPSRC Established Career Fellow. The authors acknowledge the use of the Computational Shared Facility and the Protein Structure Facility at The University of Manchester.

Author Contributions M.O. carried out molecular biology, biophysical and structural biology studies on the *P. mendocina* HODM. M.O. and L.D. carried out purification and characterisation of variant haem domain forms. P.L., B.M. and T.T. were involved in initial screening and solution characterization of HODM homologues. T.T. carried out *in vivo* formaldehyde detection. K.F. and S.E.J.R. performed and analysed EPR experiments. S.H. performed DFT calculations and kinetic data analysis. All authors discussed the results and participated in writing the manuscript. D.L. initiated and directed this research.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.L. (david.leys@manchester.ac.uk).

Reviewer Information Nature thanks P. Ortiz de Montellano, C. Wilmot and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Materials. All materials were obtained from Sigma-Aldrich unless otherwise stated.

Cloning, mutagenesis and heterologous expression in *E. coli*. For expression of an active HODM module, the gene encoding the haem α subunit (Pmen_3455 in the *Pseudomonas mendocina ymp* genome) and δ subunits (Pmen_3453) were synthesized and co-expressed from the pCOLADuet-1 plasmid (Novagen). The γ (Pmen_3456; glycine cleavage T protein – aminomethyl transferase) and β subunits (Pmen_3454; ferredoxin) were synthesized and expressed from the pETDuet-1 plasmid for a two-plasmid expression system for all four subunits (α , β , γ and δ). The α subunit bears a C-terminal His₆ tag, the δ subunit bears an N-terminal His₆ tag, whereas the γ and β subunits are untagged. For haem domain isolation, the α subunit was expressed from the pET28a plasmid with a C-terminal His₆ tag (Novagen). Site-directed mutagenesis of the α subunit within the pCOLADUET-1 vector was performed to include a Trp180Ala mutation (538–540 bp, TGG to GCG), a Arg224Ala mutation (670–672 bp, CGT to GCT) and a Glu266Gln mutation (796–678 bp, GAA to CAA). HODM variants were expressed in *E. coli* BL21 (DE3) cells (Novagen). *E. coli* transformants were grown in 1 l 2YT supplemented with 25 $\mu\text{g ml}^{-1}$ kanamycin and 100 $\mu\text{g ml}^{-1}$ ampicillin and incubated at 37 °C 180 r.p.m. until the culture reached an OD₆₀₀ of 0.6. Protein expression was then induced with addition of 0.5 mM IPTG. Cultures were grown overnight at 30 °C, 180 r.p.m. and then harvested by centrifugation (7,000g for 10 min). Cell pellets were resuspended in buffer A (100 mM Tris, 200 mM KCl pH 7.5) supplemented with DNase, lysozyme (Sigma-Aldrich) and a complete EDTA-free protease inhibitor cocktail tablet (Roche). Cells were lysed using a French press at 1,500 p.s.i. and the lysate clarified by centrifugation at 125,000g for 90 min. The supernatant was applied to a Ni-NTA agarose column (Qiagen). The column was washed with three column volumes of buffer A supplemented with 10 mM imidazole and the protein was eluted in 1 ml fractions with buffer A supplemented with 300 mM imidazole. Imidazole was removed using a 10-DG desalting column (Bio-Rad) equilibrated with buffer A. The HODM haem domain was further purified using Q-Sepharose resin at pH 8 (GE Healthcare) with a 100–500 mM elution gradient (R_z \approx 1.2). For crystallization studies, the purified protein was subjected to size-exclusion chromatography using a Superdex 200 column (GE Healthcare) in 50 mM KPi, 250 mM KCl (pH 7.5). Protein was concentrated as required using a Vivaspinn centrifugal device (Sartorius).

Haemoprotein concentration estimations. The pyridine haemochromogen method was used to quantify HODM haem and to determine an extinction coefficient at the Soret maximum ($\epsilon_{423} = 94 \text{ mM}^{-1} \text{ cm}^{-1}$) according to the method of Berry and Trumpower³¹.

UV-visible absorbance studies. UV-visible absorbance analysis was carried out using a Cary UV-50 UV-visible scanning spectrophotometer (Varian) using a 1 cm path length quartz cuvette, recording spectra between 250 and 700 nm, and typically with HODM at 8 μM in buffer A. Potassium ferricyanide was used to chemically oxidize HODM and O₂ binding experiments were performed in an anaerobic glove box (Belle Technology, Weymouth, UK) under a nitrogen atmosphere with O₂ levels maintained at less than 2 p.p.m. All solutions were deoxygenated by sparging with nitrogen gas. To obtain the ferrous dioxygen complex, HODM (holoenzyme or the haem subunit in isolation) was reduced by sodium dithionite, followed by slow bubbling of air into the reaction mixture. The rate of decay of the ferrous–O₂ complex was determined by continual scanning every \sim 2 min over \sim 6 hours. Half-life values were derived by plotting the induced optical change against time and fitting using an exponential decay equation using Origin software (OriginLab, Northampton, MA). Rate of decay values were derived using a first order rate equation.

Ligand binding titrations. Optical titrations to determine K_d values for DMA were performed at 25 °C using HODM at \sim 8 μM in buffer A. DMA stocks were made up in the same buffer and added volumes did not exceed 0.1% of the total volume. K_d values were determined by plotting the induced optical change against ligand concentration and by fitting using a hyperbolic equation within Origin software, or (for tight-binding ligands) using the Morrison (quadratic) equation³² (OriginLab, Northampton, MA).

NADPH spectrophotometric activity assays. At each concentration of DMA (0–600 μM), 200 μM NADPH was added to initiate the reaction and the rate of change in absorbance at 340 nm monitored over 1 min at room temperature to follow substrate-dependent NADPH oxidation. Reactions were performed in triplicate to produce a mean rate calculated as moles of NADPH oxidized per mol HODM per second. These data were plotted against the relevant DMA concentration and fitted using the Michaelis–Menten hyperbolic function within Origin software. Similar experiments were conducted using perdeuterated DMA or in D₂O, revealing no significant differences in NADPH oxidation rates.

***In vitro* formaldehyde detection using Purpald.** For the detection of substrate-dependent formaldehyde production, the Zurek and Karst method³³ was used, however with a Purpald solution of 1 mg ml^{−1} in 1 M NaOH. HODM activity reactions were set up as described above with an enzyme concentration of 200 nM with 250 μM DMA and allowed to react for 5 min at room temperature. Microtitre plate assays were performed using a 96-well plate reader (BioTek Synergy HT, Biotek). 100 μl of Purpald solution was mixed with 100 μl of sample and absorbance measurements were carried out after a reaction time of 60 min at a wavelength of 550 nm. A formaldehyde standard curve was calculated at known concentrations of formaldehyde. Reactions without HODM or NADPH result in the loss of formaldehyde formation. The assay underestimated this loss owing to background absorbance from the standard curve and so data were normalized accordingly.

***In vivo* formaldehyde detection.** Enzymatically catalysed formaldehyde production was assayed *in vivo* as reported previously¹⁴. In brief, the respective assay makes use of a UV-light-sensitive GFP (excitation_{max} = 395 nm, emission_{max} = 509 nm) fused to FrmR as sensor and is hosted in *E. coli* W3GM as a formaldehyde-sensitive host. Following transformation with expression constructs of formaldehyde-producing enzymes, this reporter strain allows quantitative measurement of cytosolic formaldehyde levels *in vivo*. The corresponding assays were performed using a 96-well microtitre plate reader (BioTek Synergy HT, Biotek, USA) equipped with a 360_{Ex}/528_{Em} nm filter set with readings taken in the early stationary phase (22 to 24 h after inoculation). The data were corrected for background fluorescence of uninduced controls and adjusted for cell density and enzyme activity as necessary. All assays were done as sixfold biological replicates at least, and the results were found to be statistically significant within a confidence interval of 3%. Low levels of formaldehyde were measured colourimetrically with acetylacetone/ammonium acetate as described³⁴. Aliquots were taken (0–1 ml) and made up to a final volume of 1 ml with water and combined with 0.03 ml 7.5 M HCl and 0.25 ml of a 4 \times Nash reagent solution, prior to incubation at 38 °C for 90 min. Samples were centrifuged for 5 min at 14,000 r.p.m. in a microfuge and the A₄₁₄ determined in a 1-cm cell.

Electron paramagnetic resonance (EPR) analysis. Continuous wave EPR spectra for HODM were recorded at X-band (\sim 9.4 GHz) using a Bruker ELEXSYS E500/E580 EPR spectrometer (Bruker GmbH, Rheinstetten, Germany). Temperature was maintained using an Oxford Instruments ESR900 helium flow cryostat coupled to an ITC 503 controller from the same manufacturer. The experimental conditions employed to obtain individual spectra are given in the figure captions. EPR sample tubes were 4 mm Suprasil quartz supplied by Wilmad (Vineland, NJ).

Redox potentiometry. To determine the midpoint potential for the HODM haem domain Fe^{III}/Fe^{II} couple, a spectroelectrochemical titration was performed in an anaerobic glove box (Belle Technology, Weymouth, UK) under a nitrogen gas atmosphere. All solutions were degassed under vacuum with nitrogen prior to use in the glove box. Oxygen levels were maintained at less than 5 p.p.m. The concentrated HODM haem domain protein sample was passed through a Sephadex G25 column (1 \times 20 cm) (10DG column, Bio-Rad, Hemel Hempstead, UK) immediately on admission to the glove box to remove all traces of oxygen. This column was pre-equilibrated and proteins were buffer-exchanged into anaerobic redox buffer (100 mM Tris, 200 mM KCl, pH 7.0). The HODM haem domain solution (\sim 10–15 μM enzyme in 5 ml buffer), was titrated electrochemically according to the method of Dutton³⁵, using sodium dithionite as reductant and potassium ferricyanide as oxidant. Mediators were added to expedite electronic equilibration in the system (2 μM phenazine methosulfate [PMS], 5 μM 2-hydroxy-1,4-naphthoquinone [HNQ], 5 μM methyl viologen [MV], and 1 μM benzyl viologen [BV]) and to mediate in the range between approximately +100 mV and −480 mV versus the normal hydrogen electrode (NHE)³⁶. 10–15 minutes were allowed to elapse between each addition of reductant/oxidant to ensure equilibration and stabilization of the electrode. Spectra (250–750 nm) were recorded using a Cary UV-50 Bio UV-Visible scanning spectrophotometer coupled to a fibre optic probe immersed in the HODM haem domain solution. The potential was measured using a SevenEasy S20-K meter (Mettler Toledo, Leicester, UK) coupled to a Calomel electrode (ThermoRussell, Cupar, UK) at 25 °C. The calibration of the electrode was done by using the Fe^{III}/Fe^{II} EDTA couple as a standard (+108 mV). The electrode reading was corrected by +244 mV relative to the NHE. Absorbance data at wavelengths reporting on the transition of the haem Soret band between oxidized (Fe^{III}) and reduced (Fe^{II}) forms of the HODM haem domain were plotted versus applied potential, and the data were fitted using the Nernst function in Origin 8.5 (OriginLab, Northampton MA).

Stopped-flow spectroscopy. Stopped-flow absorbance measurements were made using an Applied Photophysics SX18 MVR stopped-flow spectrophotometer in an anaerobic glove box (Belle Technology Weymouth, UK). The drive syringes were loaded with separate solutions of ferric or ferrous-oxy protein (generation of this complex with aerobic buffer is described above) and varying concentrations of reductant in buffer A in the presence and absence of DMA (260 μM). Multiple wavelength data were collected at room temperature using a photodiode

array (PDA) detector and XSCAN software. The resulting spectra were subjected to global analysis using SpecFit/32 software (Bio-Logic Science Instruments, Grenoble, France). Singular value decomposition of the dataset provides an assessment of the number of spectrally distinct components present in the reaction. The data were fitted using irreversible models.

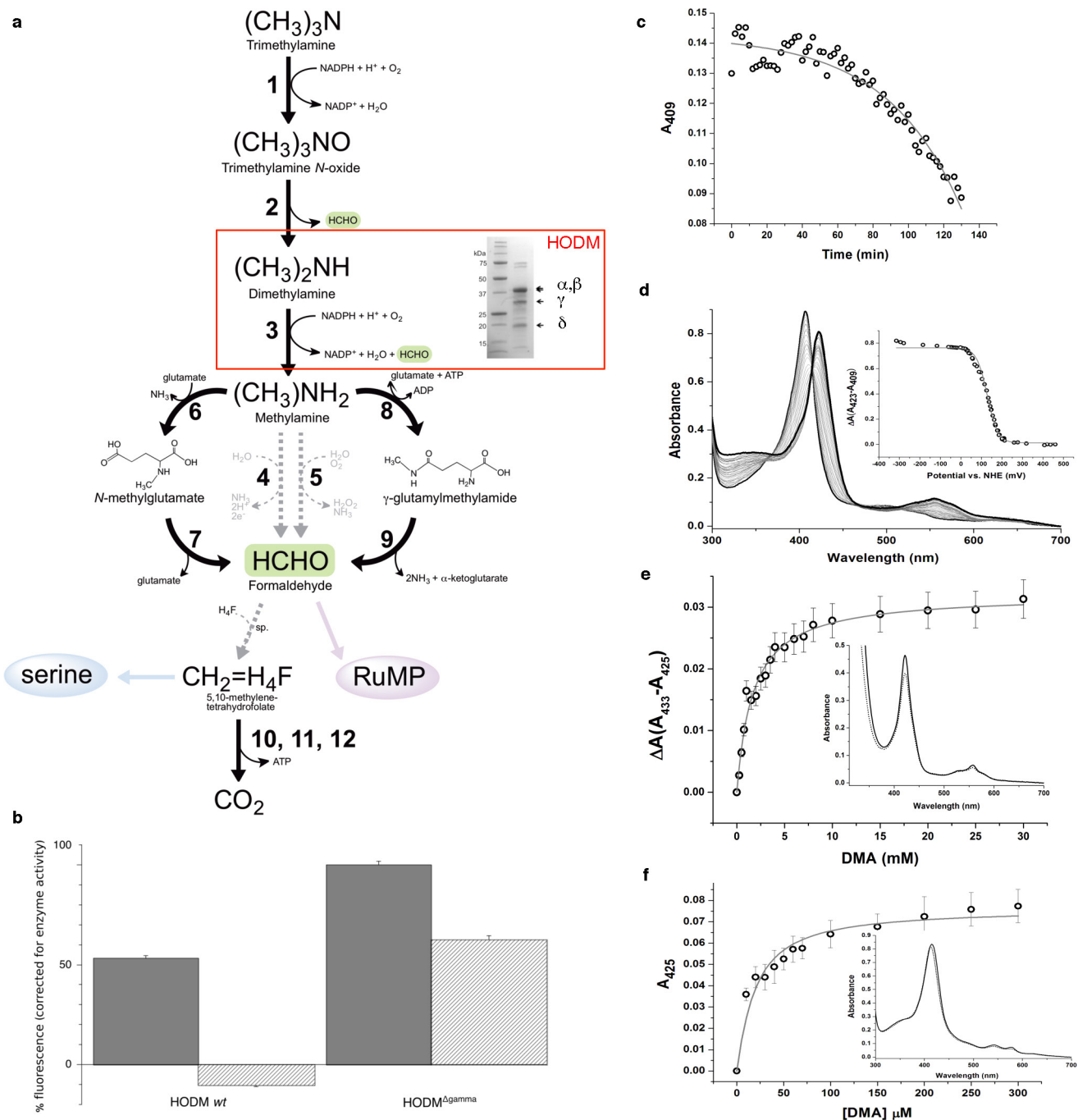
Crystallization, refinement and model building. The HODM haem subunit was crystallized at 3 mg ml⁻¹ in 50 mM KPi, 100 mM KCl, pH 7.5. Initial crystallization conditions were identified using the JCSG+ matrix screen (Molecular dimensions). Crystals suitable for diffraction experiments were obtained by sitting drop vapour diffusion at 4 °C in 400 nl drops containing equal volumes of protein and a solution containing 30% PEG 2K MME and 0.1 M potassium thiocyanate. The crystals were derivatized by soaking with the same solution supplemented with 1 mM mercury (II) acetate for 10 min. For the DMA-bound structures, native crystals were soaked with 1 mM ligand for 10 min. For ferrous structures, the crystals were soaked with sodium dithionite followed by 1 mM NaNO₃ and 1 mM DMA for ferrous–DMA–nitric oxide bound structures. The crystals were cryoprotected by the addition of 10% PEG 200 to the mother liquor and flash-cooled in liquid nitrogen. Data were collected on beamline I24 (wavelength 0.9173/1.0118 Å) at the Diamond Light Source Facility and reduced and scaled with the X-ray Detector Software suite (XDS³⁷). The ligand-free HODM haem subunit crystal structure was determined by SIRAS using the MLPHARE program in the CCP4 suite³⁸. In brief, the SIRAS signal extending to 1.86 Å from the comparison of a platinum derivative with a mercury derivative with a figure of merit of 0.85, phasing power of 1.04^{centric}/1.19^{acentric} and R_{cullis} of 0.7^{centric}/0.8^{acentric} and 1^{anomalous} was used for initial map calculation following density modification using the DM program in the CCP4 suite. A non-isomorphous native dataset with higher resolution (1.65 Å) was used for refinement. Automated model building with BUCCANEER in the CCP4 suite successfully constructed the entire model of the ligand-free HODM haem subunit structure. The model was completed by iterative cycles of manual

model building and real space refinement using the program Coot and crystallographic refinement using Refmac5. The processing and final refinement statistics are presented in Extended Data Table 1.

Single-crystal microspectrophotometry. Changes in visible absorbance spectra during X-ray exposure were monitored at the European Synchrotron Radiation Facility (ESRF) in Grenoble on beamline ID14-1 using an on-line spectrophotometer (OCEAN OPTICS DH 2000 light source and HR 2000 detector). A HODM α subunit crystal was soaked with sodium dithionite and DMA followed by transfer to mother liquor supplemented with DMA in the absence of dithionite to allow for oxygen complex formation. Initial spectra prior to X-ray illumination revealed formation of a species consistent with the ferrous–O₂–DMA complex.

Data availability. Coordinates and structure factors have been deposited in the Protein Data Bank under accession numbers 5LTE, 5LTH and 5LTI.

31. Berry, E. A. & Trumpower, B. L. Simultaneous determination of hemes a, b, and c from pyridine hemochrome spectra. *Anal. Biochem.* **161**, 1–15 (1987).
32. Morrison, J. F. Kinetics of the reversible inhibition of enzyme-catalysed reactions by tight-binding inhibitors. *Biochim. Biophys. Acta* **185**, 269–286 (1969).
33. Zurek, G. & Karst, U. Microplate photometric determination of aldehydes in disinfectant solutions. *Anal. Chim. Acta* **351**, 247–257 (1997).
34. Nash, T. The colorimetric estimation of formaldehyde by means of the Hantzsch reaction. *Biochem. J.* **55**, 416–421 (1953).
35. Dutton, P. L. Redox potentiometry: determination of midpoint potentials of oxidation-reduction components of biological electron-transfer systems. *Methods Enzymol.* **54**, 411–435 (1978).
36. Munro, A. W., Noble, M. A., Robledo, L., Daff, S. N. & Chapman, S. K. Determination of the redox properties of human NADPH-cytochrome P450 reductase. *Biochemistry* **40**, 1956–1963 (2001).
37. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
38. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).

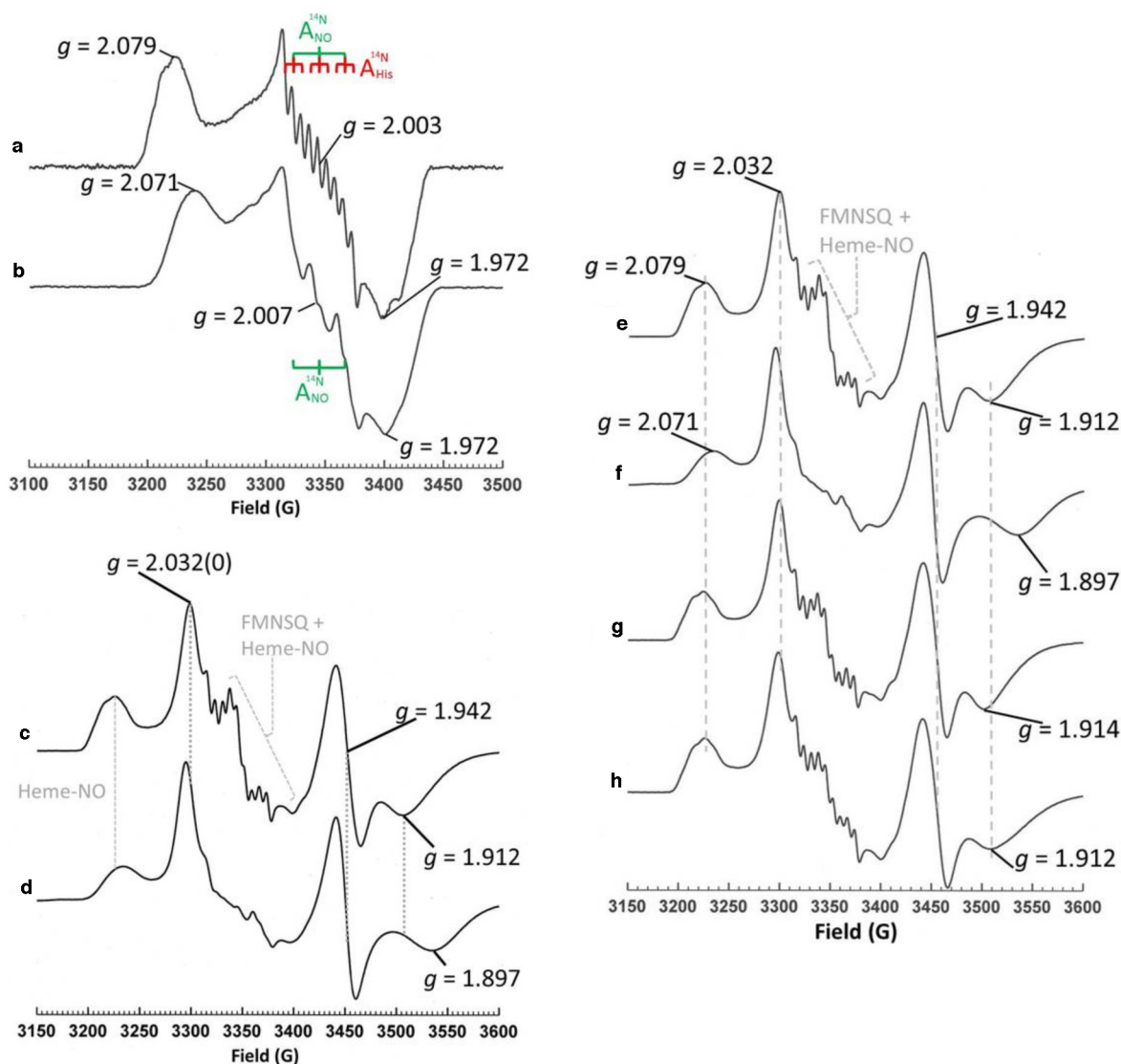


Extended Data Figure 1 | See next page for caption.

Extended Data Figure 1 | Catabolic pathway featuring HODM and additional HODM characterization *in vivo* and *in vitro*.

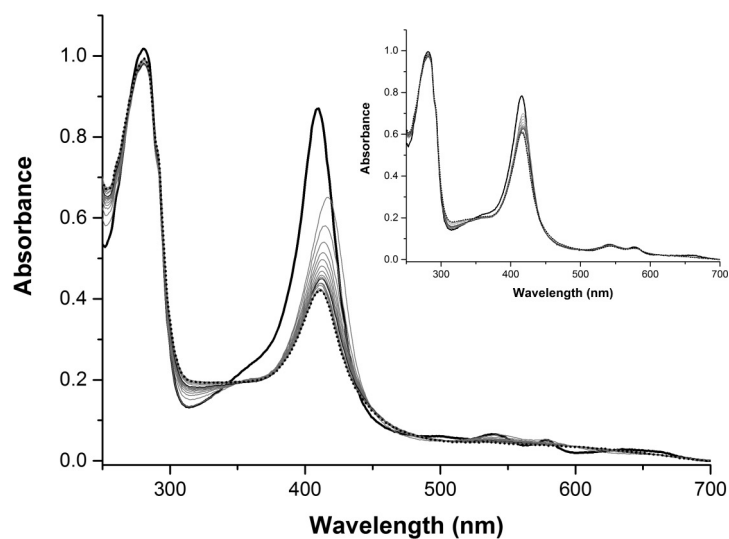
a, Trimethylamine degradation pathway with links to C1 metabolism in proteobacteria. 1, trimethylamine monooxygenase; 2, trimethylamine *N*-oxide demethylase; 3, haem-dependent oxidative demethylase (HODM), also known as secondary amine monooxygenase (SAMO); 4, methylamine dehydrogenase (EC 1.4.99.3); 5, methylamine oxidase (EC 1.4.3.21); 6, *N*-methylglutamate synthase (EC 2.1.1.21); 7, *N*-methylglutamate dehydrogenase (EC 1.5.99.5); 8, γ -glutamylmethylamide synthetase (EC 6.3.4.12); 9, γ -glutamylmethylamide-dissimilating enzyme; 10, methylene-H4F dehydrogenase/cyclohydrolase (EC 1.5.1.15); 11, formyl-H4F deformylase/ formyl-H4F synthetase (EC 3.5.1.10); 12, formate dehydrogenase (EC 1.2.1.2). Formaldehyde can be assimilated via the ribulose monophosphate (RuMP) pathway or, via methylene-tetrahydrofolate, through formation of serine. The inset shows an SDS-PAGE gel of purified recombinant HODM, revealing bands corresponding to all four subunits. Additional bands are visible that appear to correspond to multimeric forms or proteolytic fragments of HODM subunits as verified by mass spectrometry of tryptic digests. **b**, Formaldehyde leakage from HODM *in vivo*. Efficient detoxification of formaldehyde requires the γ subunit as well as tetrahydrofolic acid (THF). Cells containing wild-type and mutant HODM enzyme ($\Delta\gamma$, deletion of subunit gamma) from *Methylobacillus flagellatus* (strain KT) were grown in the presence of DMA (grey bars) or DMA and glycine (shaded bars), respectively. Fluorescence readings were subsequently corrected for background activity and scaled with the mutant enzyme set as 100%. The data show that glycine, as well as

the γ subunit, have strong effects on enzymatic formaldehyde production. While the first increases intracellular levels of THF, the latter is required for THF binding. **c**, The HODM ferrous-oxy complex is long-lived at room temperature. The ferrous-oxy decay to the ferric state was monitored at 409 nm and plotted as a function of time. The decay curve observed was fitted using the exponential decay equation to derive a half-life of 50 ± 10 min with a decay rate of $k = 0.014 \pm 0.001 \text{ min}^{-1}$. **d**, Determination of the redox potential of the HODM haem subunit. The main panel shows UV-visible spectra for the HODM haem domain during a redox titration. The oxidized enzyme (thick black line, spectrum recorded at -76 mV versus normal hydrogen electrode) has its Soret maximum at 409 nm. The fully reduced enzyme (thick line, spectrum recorded at -317 mV versus NHE) has its Soret maximum at 422 nm. Intermediate spectra are indicated in thin lines, and there is an isosbestic point at approximately 415 nm. The HODM haem domain also displays increased absorbance in the Q-band region (~ 520 – 620 nm) in the reduced (ferrous) state. The inset shows a plot of absorbance change (ΔA_{422} minus ΔA_{409}) versus applied potential (versus NHE) fitted using the Nernst equation. This provides a midpoint reduction potential value for the HODM haem domain $\text{Fe}^{\text{III}}/\text{Fe}^{\text{II}}$ couple of 128 ± 4 mV versus NHE. **e**, **f**, DMA binding to ferrous HODM (**e**) and DMA binding to ferrous-oxy HODM (**f**). Absorbance changes (as shown in the inset) as a function of DMA concentration are fitted using a hyperbolic function/Morrison equation, leading to a $K_d = 1.7 \pm 0.2$ mM for ferrous HODM (ferric HODM has a K_d of 1.5 ± 0.1 mM, data not shown) and a $K_d = 15 \pm 3$ μM for ferrous-oxy HODM (errors bars are s.e.m., $n = 3$).

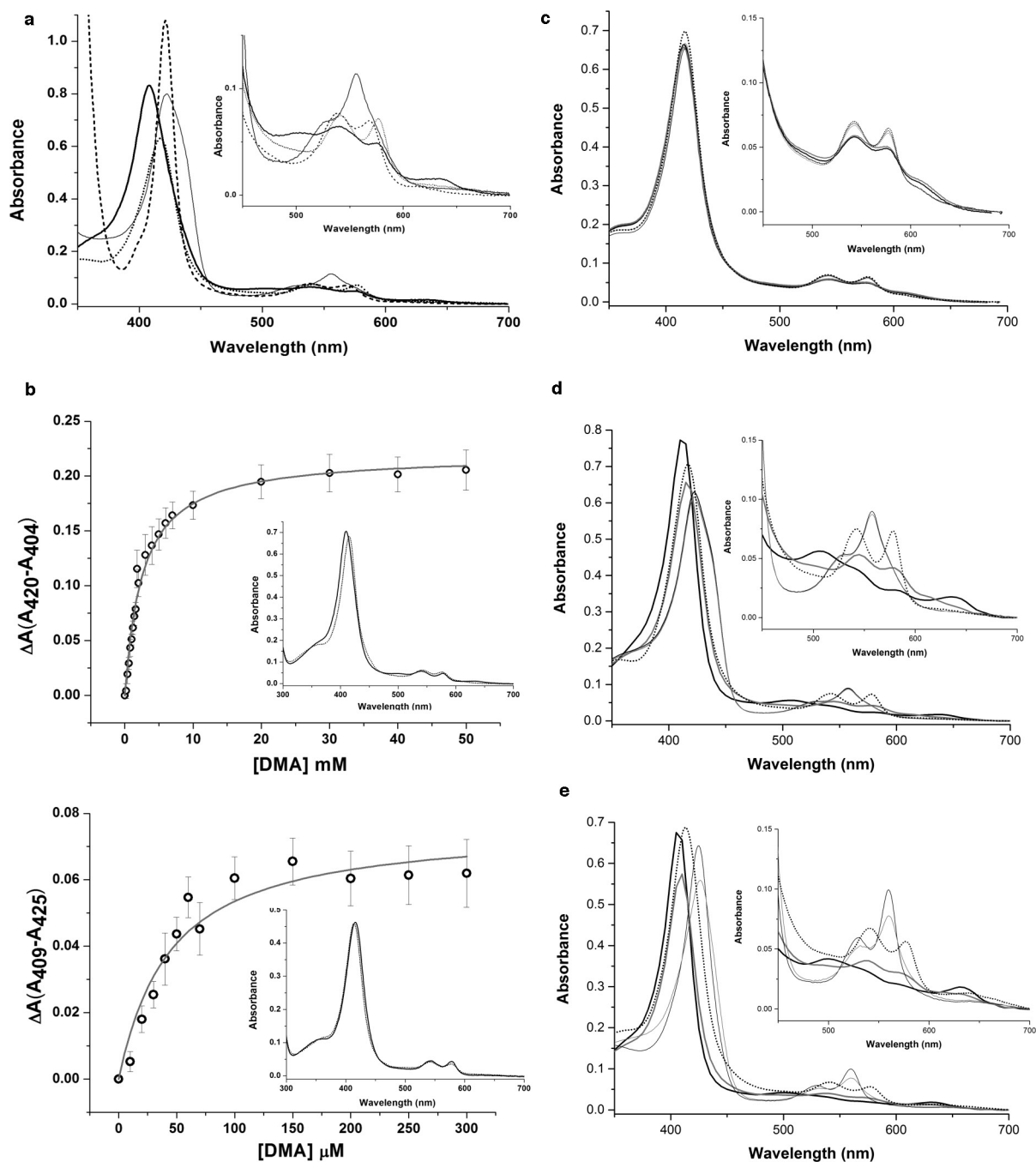


Extended Data Figure 2 | X-band continuous-wave electron paramagnetic resonance spectra of anaerobically purified HODM. **a**, Anaerobically purified HODM exposed to air for two minutes. **b**, The HODM enzyme shown in **a** with the addition of excess DMA. **c**, Anaerobically purified HODM. **d**, The HODM enzyme shown in **c** with the addition of excess DMA. Experimental parameters: microwave power, 0.5 mW; modulation amplitude, 3 G; temperature, 30 K. **e**, HODM alone.

f, HODM plus excess DMA. **g**, HODM plus excess diethylamine. **h**, HODM plus excess piperidine. Experimental parameters: microwave power, 0.5 mW; modulation amplitude, 3 G; temperature, 30 K. Only DMA (**f**) shows the loss of the ^{14}N His superhyperfine coupling in the ferrohaem-NO component of the spectrum and a significant change in the g values exhibited by the $[2\text{Fe-2S}]^{1+}$ component of the spectrum.

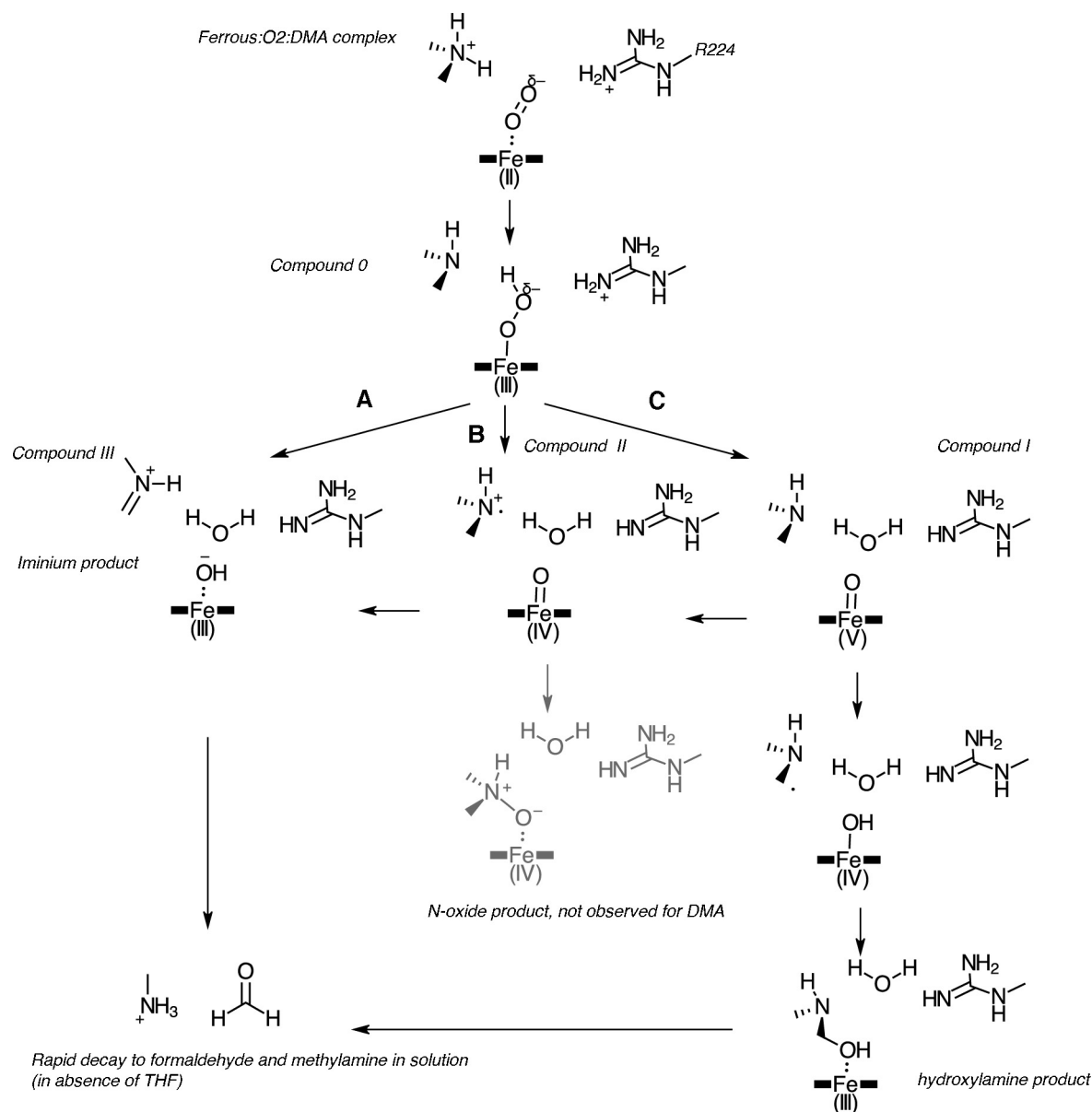


Extended Data Figure 3 | HODM cannot make use of peroxide. Spectral changes observed following mixing of $10\ \mu\text{M}$ ferric HODM with $200\ \mu\text{M}$ H_2O_2 . The haem Soret peak (black line) decays upon H_2O_2 addition ($200\ \mu\text{M}$) (thin grey lines each after 30 s) with further decay at 5 min incubation (thin black line) and 60 min (black dotted line). The inset shows a similar experiment conducted in the presence of 2 mM DMA.



Extended Data Figure 4 | The HODM haem subunit has similar properties to the HODM holoenzyme. a, UV-visible absorption features of alpha HODM (8.5 μ M). Ferric (thick solid line), ferrous (thin solid line) $\text{Fe}^{2+}-\text{O}_2$ (dotted line) and $\text{Fe}^{2+}-\text{CO}$ (dashed line). The major (Soret) absorption band is centred at 409, 423, 418 and 421 nm, respectively. The α/β bands are magnified in the inset. **b**, Titration of ferric and ferrous-oxy HODM α subunit with DMA. Absorbance changes associated with the haem Soret peak are plotted as a function of DMA concentration. The data are fitted using a hyperbolic function or the Morrison equation. Top, ferric with $K_d = 12.6 \pm 0.2$ mM. Bottom, ferrous-oxy with $K_d = 41.0 \pm 5.0$ μ M (errors bars are s.e.m., $n = 3$). Insets show UV-vis spectra of the titration of DMA (0–50 mM) against ferric and ferrous-oxy HODM α subunit

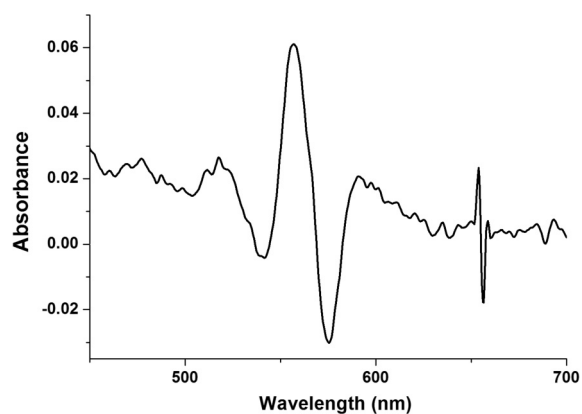
(7.5/4.8 μ M). The UV-vis spectrum of α in buffer A was recorded initially and then after each addition of DMA. The direction of the absorbance changes of the ferric and ferrous-oxy Soret peak on substrate addition is measured from 409/415 nm (thick line) to 416/414 nm (dotted line), respectively. **c–e**, Spectral properties of HODM haem subunit variants E266Q, W180A and R224A, respectively. Ferric (thick black line), ferric in presence of 1 mM DMA (thick grey line), ferrous (thin black line), ferrous in presence of 1 mM DMA (thin grey line) and $\text{Fe}^{2+}-\text{O}_2$ (dotted line). The α/β bands are magnified in the inset. While the E226Q variant lacks spectroscopic features associated with DMA binding, the R224A variant appears to bind DMA as a 6th ligand to the haem iron in the ferrous state. Oxygen binding can be detected for all variants.



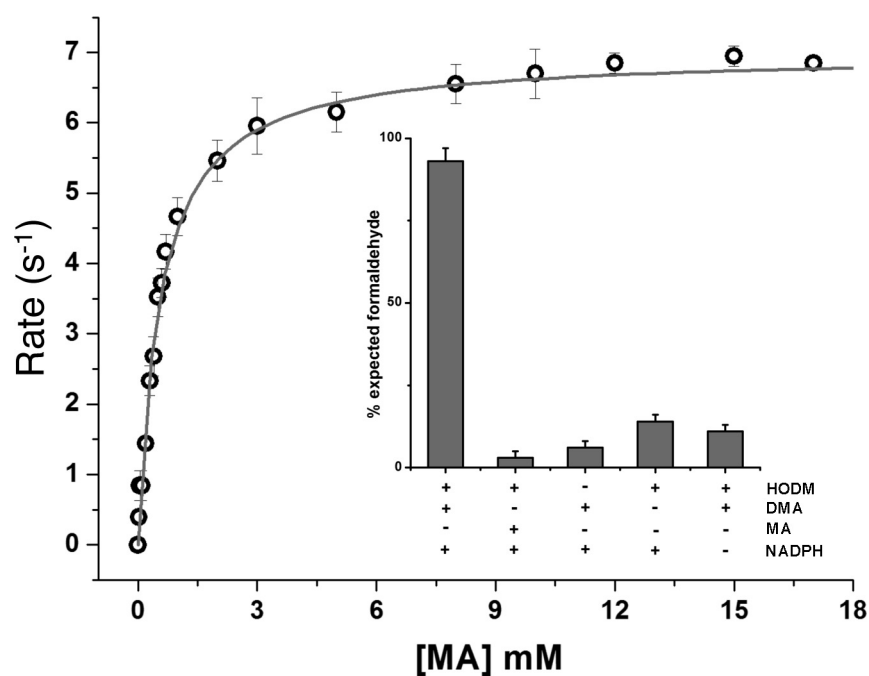
Extended Data Figure 5 | Possible routes for DMA oxidation in HODM.

Following formation of compound 0, various possibilities exist for O–O bond cleavage in principle. A direct hydride transfer from the substrate is shown in A, bypassing the need for higher oxidation states of the haem Fe. Route B depicts the homolytic bond cleavage with formation of compound II and an amine radical. Route C depicts a P450-like

mechanism, with formation of the highly oxidizing compound I. Both the B and C routes can give rise to various products. The N-oxide product is not observed with DMA as a substrate, which results in stoichiometric production of formaldehyde *in vitro*. This can be derived from either the iminium or hydroxylamine products.

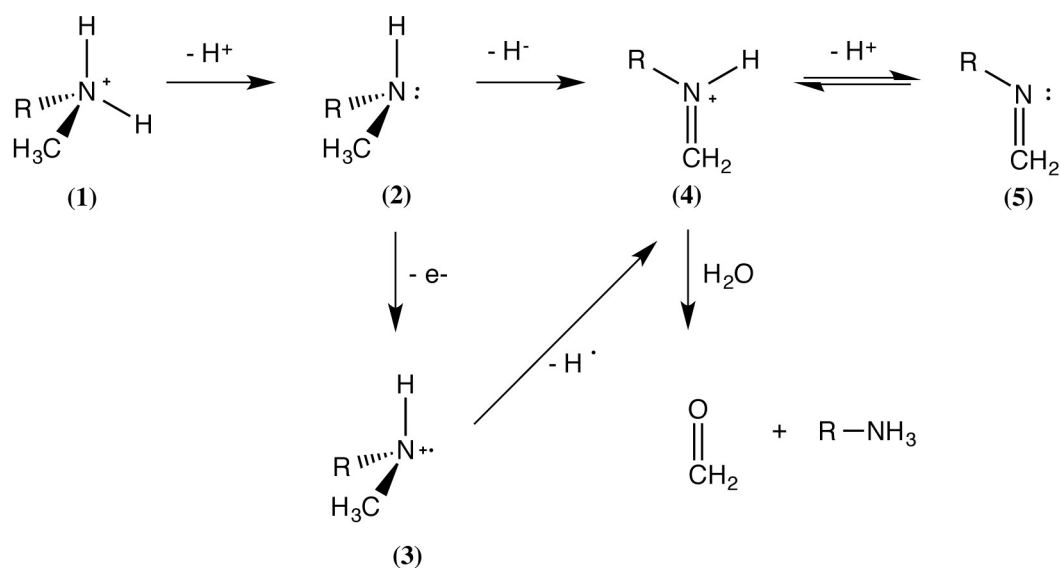


Extended Data Figure 6 | X-ray-induced photoreduction of the HODM haem subunit in complex with DMA and O₂. The initial spectrum recorded before X-ray illumination (ESRF beamline ID14-1 (wavelength 0.934 Å)) was subtracted from the post illumination spectrum. A shift in the α/β bands is recorded with a decrease at 575 nm and an increase at 556 nm. This corresponds to formation of the ferrous species (merged α/β band at 556 nm).



Extended Data Figure 7 | MA-dependent NADPH oxidation by HODM in absence of formaldehyde production. The kinetic curve (mean values of initial rate plotted as a function of MA concentration) was fitted using the Michaelis–Menten function $k_{\text{cat}} = 7.0 \pm 0.1 \text{ s}^{-1}$ and

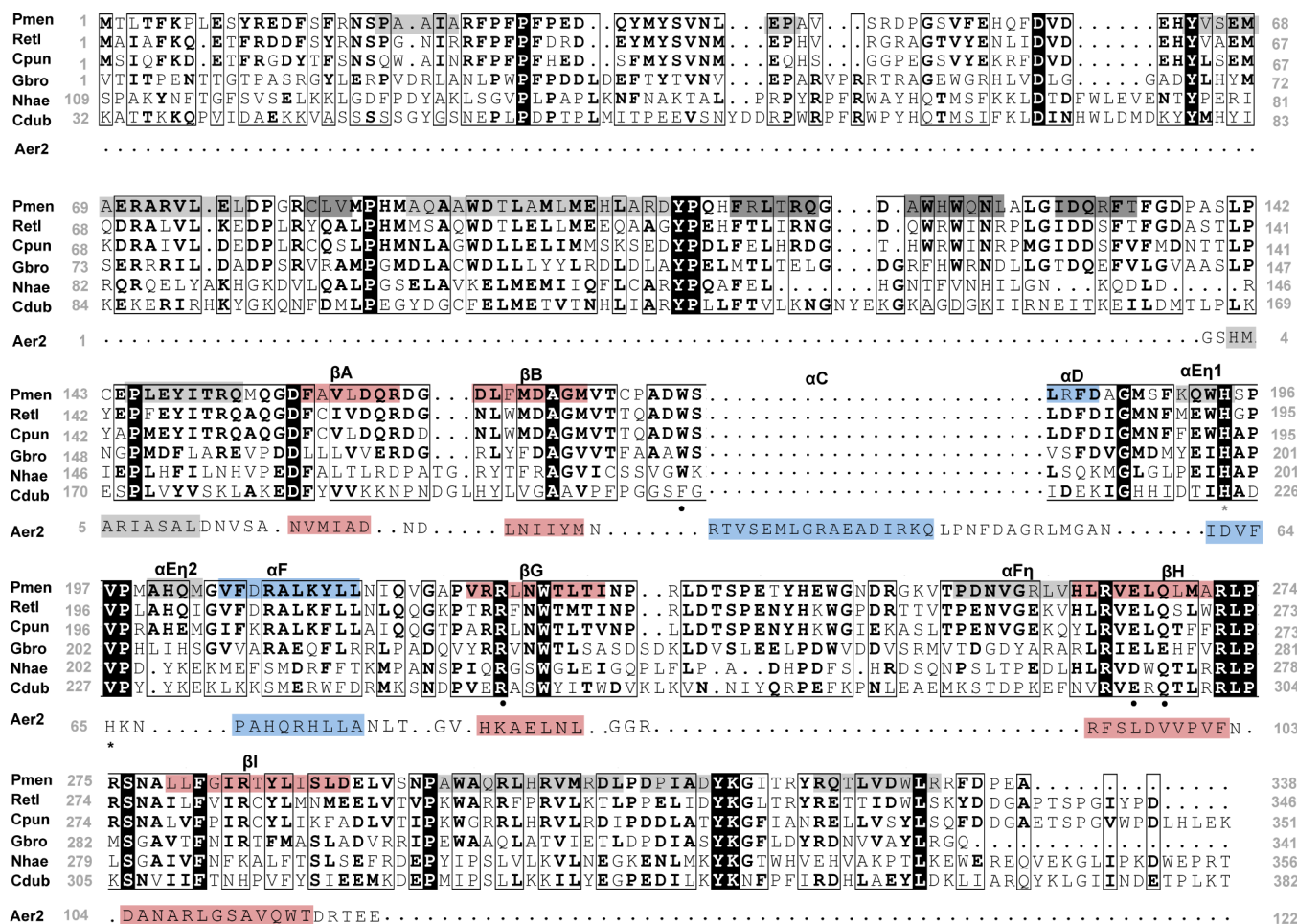
$K_M = 0.55 \pm 0.04 \text{ mM}$. Inset shows Purpald colourimetric assay for detection of substrate-dependent formaldehyde production. Data are represented as a percentage of expected formaldehyde produced for 1:1 conversion of DMA/MA to formaldehyde. Error bars represent s.d.



	(U)MP2		(U)B3LYP	
	dimethylamine	methylamine	dimethylamine	methylamine
(2) neutral amine	-134.67861 (0.00000)	-95.52167 (0.00000)	-135.10998 (0.00000)	-95.82237 (0.00000)
(3) radical cation	-134.33122 (0.34739)	-95.15169 (0.36998)	-134.76707 (0.34291)	-95.45484 (0.36753)
(4) protonated imine	-133.83615 (0.84246)	-94.66587 (0.85580)	-134.25113 (0.85884)	-94.94862 (0.87375)
(5) deprotonated imine	-133.49597 (1.18263)	-94.33499 (1.18669)	-133.90958 (1.20040)	-94.61696 (1.20542)

Extended Data Figure 8 | Gas phase density functional theory and *ab initio* comparison of the energies of DMA- and MA-based reaction intermediates. Quoted energies in the table are the sum of electronic and thermal enthalpies for structures in the gas phase at 298.15 K with units of hartrees. Values in parentheses are the difference relative to the neutral amine. The relative energies of the DMA species are lower than those of

the MA species by 10–65 kJ mol^{-1} (0.04–0.025 hartrees). The pK_a values of DMA ($\text{R} = \text{CH}_3$) and MA ($\text{R} = \text{H}$) (1) are very similar (10.64 and 10.62 respectively). Calculations were performed at both the (U)B3LYP/6-311++G(d,p) and (U)MP2/6-311++G(d,p) level using the 'Freq' keyword in Gaussian 09 rev. B.01.



Extended Data Figure 9 | Multiple sequence alignment of HODM including Aer2 PAS domain secondary structure information. Alignment of HODM proteins from the DUF3445 family from the bacteria: *Pseudomonas mendocina* (Pmen), *Rhizobium etli* (Retl), *Candidatus Punicispirillum marinum* (Cpun) and *Gordonia bronchialis* (Gbro), and from the fungi *Candida dubliniensis* (Cdub) and *Nectria*

haematococca (Nhae). The catalytic Arg, Glu and Gln residues and the substrate-discriminating Trp are marked with black dots. The conserved haem-binding His in Pmen and Aer2 are marked with a grey and a black asterisk, respectively. Secondary structural elements are highlighted, PAS-domain-specific α -helices in blue and β -sheets in pink, and HODM-specific α -helices in pale grey and β -sheets in dark grey.

Extended Data Table 1 | Data collection and refinement statistics for *P. mendocina* HODM α subunit

<u>HODMalpha</u>	Ferric	Ferric-DMA	Ferrous-DMA-NO
PDB code	5LTE	5LTH	5LTI
Data collection			
Space group	P 4 ₃ 2 ₁ 2	P 4 ₃ 2 ₁ 2	P 4 ₃ 2 ₁ 2
Cell dimensions			
<i>a</i> , <i>c</i> (Å)	79.91, 144.71	80.38, 144.79	80.21, 144.61
Resolution (Å)	26.82-1.65 (1.70-1.65)	56.84-1.76 (1.81-1.76)	28.36-1.90 (2.01-1.90)
<i>R</i> _{rim}	0.034 (0.223)	0.027 (0.211)	0.057 (0.306)
<i>I</i> / σ <i>I</i>	15.4 (3.8)	18.2 (3.5)	13.7 (4.5)
Completeness (%)	99.93 (100)	99.91 (100)	100 (100)
Redundancy	12.1 (12.2)	12.9 (12.2)	12.0 (12.0)
Refinement			
Resolution (Å)	26.82-1.65 (1.70-1.65)	56.84-1.76 (1.81-1.76)	28.36-1.90 (2.01-1.90)
No. reflections	53715 (3876)	45331 (3261)	36266 (2623)
<i>R</i> _{work} / <i>R</i> _{free}	0.163/0.194 (0.198/0.248)	0.168/0.187 (0.241/0.262)	0.173/0.201 (0.21/0.28)
No. atoms	3134	3066	3004
B-factors (Å ²)	15.86	20.76	19.82
R.m.s deviations			
Bond lengths (Å)	0.024	0.028	0.025
Bond angles (°)	2.23	2.42	2.11

Values in parentheses correspond to the outer resolution shell.

CORRIGENDUM

doi:10.1038/nature19850

Corrigendum: Slowly fading super-luminous supernovae that are not pair-instability explosions

M. Nicholl, S. J. Smartt, A. Jerkstrand, C. Inserra, M. McCrum, R. Kotak, M. Fraser, D. Wright, T.-W. Chen, K. Smith, D. R. Young, S. A. Sim, S. Valenti, D. A. Howell, F. Bresolin, R. P. Kudritzki, J. L. Tonry, M. E. Huber, A. Rest, A. Pastorello, L. Tomasella, E. Cappellaro, S. Benetti, S. Mattila, E. Kankare, T. Kangas, G. Leloudas, J. Sollerman, F. Taddia, E. Berger, R. Chornock, G. Narayan, C. W. Stubbs, R. J. Foley, R. Lunnan, A. Soderberg, N. Sanders, D. Milisavljevic, R. Margutti, R. P. Kirshner, N. Elias-Rosa, A. Morales-Garoffolo, S. Taubenberger, M. T. Botticella, S. Gezari, Y. Urata, S. Rodney, A. G. Riess, D. Scolnic, W. M. Wood-Vasey, W. S. Burgett, K. Chambers, H. A. Flewelling, E. A. Magnier, N. Kaiser, N. Metcalfe, J. Morgan, P. A. Price, W. Sweeney & C. Waters

Nature **502**, 346–349 (2013); doi:10.1038/nature12569

In this Letter, we have identified an important error affecting Fig. 4 and Extended Data Fig. 6, as well as the values of some parameters derived from our model fits. We stress that this error in no way affects the discussion or the conclusions. In building the bolometric light curve of the superluminous supernova PTF 12dam, our code assumed that photometry from the Swift satellite was calibrated in the Vega magnitude system. However, our photometry was actually calibrated to the AB magnitude system (and published in the AB system in the original paper). This led to an underestimate of approximately 50% in the measured peak luminosity of PTF 12dam.

Here we present updated figures and model fits with the correct bolometric luminosity. To construct the bolometric light curve, we transformed the Swift data into Vega magnitudes, and then converted all photometry to fluxes. At epochs with the full range of UVW2 to K band, we simply integrated over the observed spectral energy distribution. At epochs with missing filters, we accounted for the unobserved flux by fitting blackbodies to the available data. We also compared our blackbody extrapolations against polynomial fits to the ultraviolet and near-infrared light curves, finding consistent results.

This should be more reliable than our previous extrapolation method, which assumed linear colour evolution over 40 days. If anything, the improved bolometric light curve strengthens our main conclusion—that PTF 12dam was not a pair-instability supernova—because the brighter light curve peak results in an even steeper rise to maximum. It is important to note that the large discrepancy compared to pair-instability models does not rely solely on bolometric comparisons: the difference was clearly apparent in the *r*-band light curves in our original

Fig. 1. Therefore this is a robust result independent of any time-varying bolometric correction.

Our secondary conclusion—that spin-down of a nascent magnetar can satisfactorily explain the observed properties—also remains intact. The parameters of our magnetar-powered fit to the corrected bolometric light curve shown in Fig. 1 of this Corrigendum (the corrected original Fig. 4) remain within a sensible range and are as follows: magnetic field $B = 5 \times 10^{13}$ G; spin period $P = 2.3$ ms; and ejecta mass $M_{\text{ej}} = 7M_{\odot}$ (where M_{\odot} is the solar mass) for an opacity $\kappa = 0.1 \text{ cm}^2 \text{ g}^{-1}$ and explosion energy $E = 10^{51}$ erg. Our suggestion that a relatively lower spin period and larger ejected mass can explain the existence of these long-duration superluminous supernovae is unchanged.

In Extended Data Fig. 6, we showed that decay of radioactive nickel-56 could not explain the observed light curve. This remains true for the corrected light curve, which is shown in the Supplementary Information to the Corrigendum. The unrealistic parameters required to model the data with nickel as the power source are listed below the figure.

We thank P. Vreeswijk for initially pointing out a discrepancy between our light curve and his own results. M.N. identified the source of the discrepancy. The original Letter has not been corrected.

Supplementary Information is available in the online version of the Corrigendum.

Correspondence and requests for materials should be addressed to M.N. (matt.nicholl@cfa.harvard.edu).

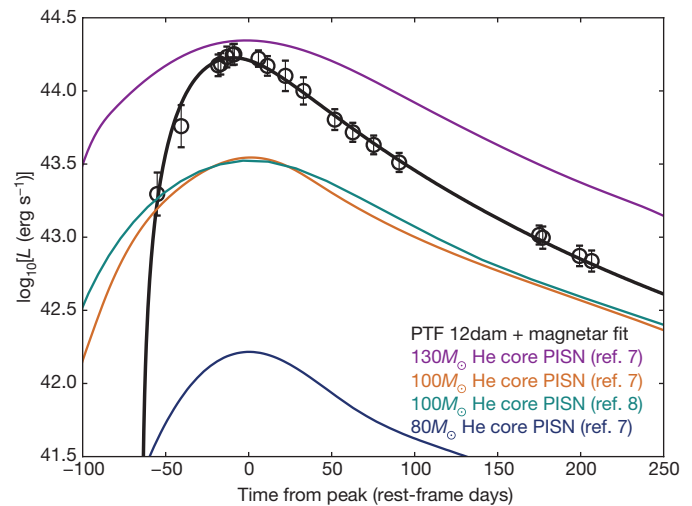


Figure 1 | This is the corrected Fig. 4 of the original Letter. PISN, pair-instability supernova. Error bars are 1σ , as in the original Letter.

CAREERS

BLOG Counsel and stories to help you reach career goals blogs.nature.com/naturejobs

@NATUREJOBS Follow us on Twitter for news and features go.nature.com/e492gf

NATUREJOBS For the latest career listings and advice www.naturejobs.com

ALEX GARLAND



Former climate scientist Jess Spear (third from left), now an organizer for the activist group Socialist Alternative, campaigns on the streets of Seattle in 2014.

ACTIVISM

A call to serve

The political arena need not be off-limits to scientists.

BY CHRIS WOOLSTON

As if full-time research weren't time-consuming and challenging enough, nanophysicist Michael Stopa embraced a second occupation while at the bench: politics. He served as a delegate for US president-elect Donald Trump at this year's Republican National Convention. Before that, while he was a senior scientist at Harvard University in Cambridge, Massachusetts, he blew his cover as a semi-secret conservative by running unsuccessfully as a Republican for the US Congress in 2010 and again in 2013. "My face was on the front page of the Harvard *Crimson*," he says of the university's student newspaper. "At that point, I was exposed."

Stopa, who now works at a technology

startup near Boston, Massachusetts, says that his outspoken politics have cost him at least one close professional collaboration — and maybe more — but that hasn't quietened him. He still talks politics on the *Harvard Lunch Club* weekly podcast. In each session he takes part in, he discusses his conservative views, including his belief that illegal immigration threatens the United States.

The acrimonious US presidential election is over, but politics are forever, and Stopa isn't the only scientist joining the fray. Many researchers take public political stands on Twitter and elsewhere, and some are engaging with political parties or running for office (see 'Join the party'). Politically active scientists can struggle to find the time and energy to bridge both worlds, and there's always the risk that

an unpopular stand could cause friction. But there are also benefits: politics can provide another avenue for networking and outreach. And, ideally, scientists will be able to give governments the kind of input needed to produce informed policy. Political involvement can also create a sense of real-world accomplishment that is sometimes hard to find in the lab. "Nothing's more rewarding than combining the two passions," says David Mazzocchi-Jones, a neuroscientist at Keele University, UK, and a member of the local Labour Party.

Despite the opportunities, few scientists have reached high office in government. Frauke Petry, chairwoman of the right-wing Alternative for Germany Party, has a chemistry PhD, as does chancellor Angela Merkel. Of 535 members of the US House and Senate, ►

JOIN THE PARTY

Getting political

Want to get involved in politics? Here's how to take the plunge.

- **Step forward.** Scientists willing to volunteer their time and talent are always in demand. "If you go to any party office anywhere in the world and say that you really want to help out, they aren't going to say no," says David Mazzocchi-Jones, a neurologist at Keele University, UK, who is active in the UK Labour Party. "Or call up a local politician and offer to help. No politician will turn you away."
- **Find your issue.** Jess Spear, a former climate scientist who is now an organizer for activist group Socialist Alternative in Seattle, Washington, recommends campaigning for a local issue even if it's not specifically

science-related. If you're passionate about an increase in the minimum wage or a new tax, then making phone calls or writing letters could be a rewarding political introduction.

- **Get in the race.** Win or lose, campaigning is a proven way to spread your message. Just be ready for the verdict. "You have to have a thick skin," Mazzocchi-Jones says. "Scientists are used to rejection, but getting rejected by 40,000 people is different."
- **Join an activist group** — but only if you're ready to be a team member. "It's important that educated people come to activism with humility," Spear says. "You are there to build with others, not teach them, and you are not an expert in the political arena." **C.W.**

► just two congressmen — a physicist and an engineer — have PhDs in the hard sciences. The UK-based Campaign for Science and Engineering counts 90 Members of Parliament who have at least some background or interest in the sciences, engineering or medicine, including Thérèse Coffey, who has a PhD in chemistry. That's down from 103 science-minded MPs in the previous parliament. "Scientists are very under-represented in politics in the UK," Mazzocchi-Jones says. "Twenty years ago, there were quite a few more."

Researchers who manage to break into the political world could have a huge impact on policy, says Jeff Schweitzer, a former marine biologist who worked as a science-policy analyst for the US Clinton administration in the 1990s. "The biggest thing that a scientist brings is a method of thinking," he says. "They have a vocabulary that non-scientists might not have." Scientists in government can help to bridge the gap between policymakers and the researchers who study, in great detail, how the real world actually works, he adds.

Mazzocchi-Jones, a Labour councillor for Newcastle-under-Lyme, believes that his science background has helped him to handle the issues that matter to his constituents. "When we're deciding on a new recycling system, I can say, 'Show me the numbers,'" he says.

Governments are increasingly facing critical issues such as climate change and fracking (hydraulic fracturing) that call for scientific wisdom, says David Dunbar, a bioinformatician at the University of Edinburgh, UK, who is active with the Scottish National Party. "The scientists you see in the UK Parliament seem to be thinking in an evidence-based way, and that's a positive," he says. "The party line isn't always evidence-based. And neither is public opinion."

Scientists who aren't themselves politically active can still do their bit to keep politicians

informed, even if only through a quick e-mail or a chat with a local representative. "We need to engage more with politicians," Mazzocchi-Jones says. "It's not going to get us anywhere unless we talk to them directly." He says that his interest in politics was rekindled in 2014 when he helped to organize the UK Physiological Society's Engaging with Parliamentarians outreach programme, where he and other scientists paired up with politicians to exchange ideas. He says that both sides must find ways to identify common ground. "Scientists have to step forward and be recognized," he says, "and politicians have to listen."

Politics can be a sticky subject, however, especially when someone is out of step with their colleagues. Stopa says he felt some tension at Harvard, and not just with the friend and collaborator who severed ties with him. When Stopa's contract wasn't renewed, he was eager to move on. "It's hard to be surrounded by people with different ideologies."

Stopa doesn't regret publicly announcing his conservatism, but he understands why some conservatives prefer to keep quiet. "There's an ongoing debate about whether or not to come out of the closet," he says. "You have to make that decision for yourself. If you think it might negatively affect your career, you might be better off not saying these things."

Schweitzer sees political activism as a right. No one in science should be afraid to put their politics on display, he says. "If that's an obstacle," he adds, "you're not at the right institution."

Sometimes, political anonymity isn't much of an option. "A lot of my colleagues and students live in my ward," says Mazzocchi-Jones.

"I know for a fact that they've received leaflets with my face on them." And his dual roles occasionally collide in awkward ways. "Colleagues will tell me, 'My bin didn't get collected last week,'" he adds, by way of example.

So compelling is political work for some scientists that they turn it into a full-time profession. Stacey Danckert, who has a PhD in cognitive neuroscience from the University of Waterloo, Canada, declined a prestigious two-year grant from the Alzheimer Society of Canada in 2013 because she found it tough to balance research, politics and family commitments. "I decided to follow my passion for the environment," she says.

She left the lab and is now policy coordinator for the Green Party of Ontario and a twice-unsuccessful Green Party candidate for the Provincial Parliament of Ontario. In her view, it's almost impossible for a scientist to run for political office while staying in the lab. "It's important to get your name out, and you can't do that without spending a lot of time," she says. "The two pursuits require endless dedication."

Similarly, Jess Spear, a former climate scientist who worked at the Burke Museum of Natural History and Culture in Seattle, Washington, left research to join Socialist Alternative, a socialist activist group, in 2011. After running unsuccessfully for the Washington state House of Representatives in 2014, she is now a full-time organizer for the group. "The more I got involved in climate science," she says, "the more I became aware that we don't just need more data. We need political will."

Schweitzer believes that scientists who can handle university politics have the mettle to excel at local, regional and national politics. "The skills are very transferable," he says. "You have to show that you can get along with people, and you have to build networks."

Perhaps most importantly, scientists tend to have a track record of working with large bureaucracies. "You need to be able to manipulate the system to your will to get things done," he says. "If you tend to get frustrated and just throw up your hands, politics probably isn't for you."

Before jumping into politics, Schweitzer briefly ran his own lab at the University of California, Irvine, an experience that he says was invaluable in his second career. "In order to have credibility in Washington DC, you have to have had at least a short career as a lab scientist," he says. "I wasn't in the lab very long, but in their view I was a real scientist."

Mazzocchi-Jones manages, for the most part, to keep his work separate from his ideology. He says that he has a student who is an outspoken supporter of the UK Independence Party, a right-wing, anti-immigration party. "I find his politics abhorrent," Mazzocchi-Jones says. "But in the end, science unites us." ■

Chris Woolston is a freelance writer in Billings, Montana.

HUNGRY GHOSTS

Homing signal.

BY CASSANDRA KHAW

They keep the oldest ghosts in little diskettes, squares of ageing steel and red plastic. Except not really, of course. Human consciousness is notoriously hard to optimize. You could trim the fat in theory, Xiao Lin thinks, as she dabs carmine on the slant of her cheekbones, contouring them to knives. Delete swathes of summertime screw-ups, summarize romances, condense college to empirical scores. But then you'd lose the person in the process. You'd lose the soul.

Better to empty the diskettes and cram their corpses with solid-state drives, keep the joke and the personality. Otherwise, what's the point? Might as well skip the *getai* and head back up to the stars.

"But it's creepy, don't you think?" Her co-host this year is indistinguishable from the last: petulant, slim, hauntingly young, with pencil-stroke arms and dense, dark hair. The girl pins gold hoops to her ears and makes a moue at her reflection, earrings glittering as she tilts her chin one way and another, the light catching on silver-moted lashes. "Performing for dead people like that. We have the technology. Why don't we just schedule something for them in VR?"

Xiao Lin sighs. It's always the same. Naturalize in space and you anchor in the stars, forget what it means to be Chinese, immigrant-born and aching, your spine bent with ghosts, half-sick with longing for a place your children can call home.

"Because they can tell," she chides, not quite lying. Realistic conjecture, Xiao Lin reassures herself, smearing malachite over her eyelids. Still, who knows if the AIs can distinguish between fact and fabrication, can differentiate between polynomial renders and the jewelled sea, the saltwater breeze?

"Can they, though?"

Another layer of mascara, another coating of liner. Xiao Lin spirals her lipstick back into its casing. A resolute click.

Yes, she thinks. *Absolutely.*

But what she says is: "Do you want their families to ask for their money back?"

The invocation of money buries any counter-argument.

The two return to their preparations,

silent. Outside, the orchestra tunes its instruments, performs chords in bright arpeggios. A thrum of laughter escapes as someone reels off a joke that Xiao Lin can almost hear. And for a heartbeat, she is back



on Europa in a house teeming with relatives, the air rich with spices, a Chinese New Year comedy playing too loudly on a salvaged TV.

"I don't think they'd ask for their money back." The girl — Bao Bao or Mei Mei or something equally infantile — grumbles an interruption, timbre slightly defiant, as she secures her hair in double loops. "As long as their relatives are happy."

Xiao Lin rises, red silk frothing along the slope of her hips. She picks a *buyao* — jade peonies dangling from a filigreed twig — from the dressing table and inserts it into the girl's hair. In the dim light behind the stage, they could be sisters, features simplified into stereotype, unified by skin. "Would you be happy, though?"

"What?"

"Lying to them like that."

"To whom? The dead people?"

"Yes."

An exasperated noise. "I mean, they're dead —"

"And at some point, they weren't. They were alive and afraid of the idea that they'd be forgotten, so afraid that they made copies of themselves. Just to live on. Just to say that a part of them will always be here." Xiao Lin

works pins into her co-host's *qipao*, moulds the dress to the girl's silhouette, smaller and slighter than that of the original owner.

"I still don't understand. What's your point? Even if we just installed them into a simulation, they wouldn't know."

"But we'd know."

Clack. The sound of teeth coming together, like a coffin pulled shut.

No more argument. Only a look reflected in the mirror, bottomless, older than the girl's kohl-framed eyes, her lineless throat. Xiao Lin pats her on the shoulder before she drifts away, weaving between production assistants and extras, gofers and camera women, heels clicking a eulogy.

At the curtains, she lounges against a pillar and tweaks apart the fabric. The audience is steeped in apparitions, holograms of varying sophistication, projected onto glass and ambient moisture. Interspaced in between are crimson-clad humans, marrow and bone, middlemen facilitating long-distance conversations.

Xiao Lin winds restless fingers around the loop of gold at her throat. Twenty thousand years and a fortune of galaxies at their fingertips and still,

all they want to do is go home. So much money poured into the restoration of Earth, salving its scars, decontaminating the waters, restoring the skyline. So much effort for something they'd never see, a halcyon dream they clutched while the skies boiled with ash.

And for what? A memory carried through generations, like a song passed down from mother to daughter. A chance to let their ghosts sit and laugh and clap at performances they'd never experience, only witness by proxy. Was it worth it?

A hand finds the crook of Xiao Lin's elbow. She turns to see her co-host, grinning, iridescent with cameras. You can go anywhere, Xiao Lin muses, if you know there's somewhere to go home to. And maybe, that was the point all along.

"Ready?"

"Ready." Xiao Lin smiles and lifts the curtain, steps out to the roar of the surf and the applause of the ancestral dead. ■

Cassandra Khaw is the business developer for micropublisher Ysbryd Games. When not otherwise writing fiction, she punches things and occasionally, people.

ILLUSTRATION BY JACEY